









































[illegible]



PS Claim 1; Page 92; 113pp; English.

XX Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are

CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag

CC expression cassettes exhibit increased potency for induction of

CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV 1

CC self-assemble into non-infectious virus-like particles which are used as

CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.

XX Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;

SQ

Query Match 100.0%; Score 60; DB 21; Length 60;

Best Local Similarity 100.0%; Pred. No. 1.5e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcaagcaggcccccaaggagcccttcggcgactacgtggacgctttctcaagacc 60

|||||

Db 1 gacatcaagcaggcccccaaggagcccttcggcgactacgtggacgctttctcaagacc 60

|||||

RESULT 2

AAA51609

ID AAA51609 standard; DNA; 1479 BP.

XX AAA51609;

DT 31-OCT-2000 (first entry)

XX HIV synthetic Gag polynucleotide.

DE Gag; expression cassette; antigenic; type C, HIV, Env; synthetic;

XX DNA immunization; packaging cell line; antigen presentation; ss.

KW Human immunodeficiency virus type C strain AF110965.

XX Synthetic.

OS

XX

XX Location/Qualifiers

FT 1..1479

CDS

FT /\*tag= a

FT /product= Synthetic\_Gag

FT /note= "Codon usage pattern was modified and inhibitory

FT elements (INS) and RRE sites were inactivated

FT resulting in improved expression"

XX

XX WO200039304-A2.

XX

XX 06-JUL-2000.

XX

XX 30-DEC-1999; 99WO-US31273.

XX

XX 31-DEC-1998; 98US-0114495.

PR 01-SEP-1999; 99US-0152195.

XX

XX (CHIR ) CHIRON CORP.

XX Barnett S, Zur Megede J;

PI WPI; 2000-452401/39.

XX P-PSDB; AAY96943.

XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal

PT especially human against HIV

XX

PS Claim 2; Page 92-93; 113pp; English.

XX Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are

CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag

CC expression cassettes exhibit increased potency for induction of

CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1

CC self-assemble into non-infectious virus-like particles which are used as

CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.

XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;

SQ

Query Match 100.0%; Score 60; DB 21; Length 1479;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcaagcaggcccccaaggagcccttcggcgactacgtggacgctttctcaagacc 60

|||||

Db 844 gacatcaagcaggcccccaaggagcccttcggcgactacgtggacgctttctcaagacc 903

|||||

RESULT 3

AAA51625

ID AAA51625 standard; DNA; 1479 BP.

XX AAA51625;

AC

XX 31-OCT-2000 (first entry)

DT HIV codon-optimized synthetic Gag polynucleotide.

DE Gag; expression cassette; antigenic; type C, HIV, Env; synthetic;

XX DNA immunization; packaging cell line; antigen presentation; ss.

KW Human immunodeficiency virus type C strain AF110965.

XX Synthetic.

OS

XX WO200039304-A2.

FN

XX 06-JUL-2000.

PD

XX 30-DEC-1999; 99WO-US31273.

PF

XX 31-DEC-1998; 98US-0114495.

PR 01-SEP-1999; 99US-0152195.

XX

XX (CHIR ) CHIRON CORP.

XX Barnett S, Zur Megede J;

PI WPI; 2000-452401/39.

XX

XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal

PT especially human against HIV

XX

PS Disclosure; Page 103-104; 113pp; English.

XX

XX Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are

CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag

CC expression cassettes exhibit increased potency for induction of

CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1

CC self-assemble into non-infectious virus-like particles which are used as

CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.

XX

XX Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 other;

SQ

Query Match 100.0%; Score 60; DB 21; Length 1479;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







CC infection and acquired immunodeficiency syndrome (AIDS). The present  
 CC sequence is a common region found in Gag coding sequences: Gag protease  
 CC and Gag-polymerase  
 XX  
 SQ Sequence: 1268 BP; 273 A; 449 C; 411 G; 135 T; 0 other;  
 Query Match 89.3%; Score 53.6; DB 21; Length 1268;  
 Best Local Similarity 93.3%; Pred. No. 2.2e-08;  
 Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 qacatcaagcagagcccccagagcccttcgcgactactgagccgtttcttcaagacc 60  
 Db 862 qacatcaagcagagcccccagagcccttcgcgactactgagccgtttcttcaagacc 921  
 RESULT 9  
 AAZ52051  
 ID AAZ52051 standard; DNA: 1503 BP.  
 AC AAZ52051;  
 DT 18 JUL 2000 (first entry)  
 DE Codon optimised Human immunodeficiency virus gag coding region.  
 KW HIV; gag; packaging cell line; lentivirus; retroviral vector particle;  
 KW pol; gene therapy; gene replacement; vaccine; biochemical reagent;  
 KW codon optimisation; ds.  
 XX  
 OS Human immunodeficiency virus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1503  
 FT /\*taq d  
 FT /product= "gag protein"  
 XX  
 PN W0200015819-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 10 SEP-1999; 99WO-US20675.  
 XX  
 PR 11 SEP-1998; 98US-0100022.  
 PR 12 SEP 1998; 98US-0100063.  
 XX  
 PA (CHIL.) CHILDBRENS MEDICAL CENT.  
 XX  
 PI Gray JI. Mulligan RC;  
 XX  
 DR WP1; 2000 273455/23.  
 DR P PSB; AAY70569  
 XX  
 PS New packaging cell line for producing a viral accessory protein  
 PI independent HIV derived retroviral vector particles, useful in gene  
 PI therapy or gene replacement -  
 XX  
 PS claim 41; Fig 8; 62pp; English.  
 XX  
 CC The patent discloses new packaging cell line for producing a viral  
 CC accessory protein, independent lentiviruses, preferably  
 CC human immunodeficiency virus (HIV). Derived retroviral vector particles.  
 CC The packaging cell line comprises a mammalian cell, a retroviral DNA  
 CC comprising a coding sequence for a lentivirus, preferably HIV, gagpol,  
 CC where the coding sequence has been mutagenised to improve expression of  
 CC the viral gagpol proteins, a second retroviral nucleotide sequence  
 CC comprising the coding sequence for a heterologous envelope protein and a  
 CC third retroviral nucleotide sequence comprising a DNA sequence of  
 CC interest and lentivirus, preferably HIV, cis-acting sequences required  
 CC for packaging, reverse transcription and integration.  
 CC The packaging cell lines and viral particles can be used for gene  
 CC therapy or gene replacement with improved safety. They can also be used

CC in the development and production of vaccines and biochemical reagents.  
 CC The present sequence is a gag coding region of codon optimised HIV  
 CC gagpol sequence. This sequence is used in the packaging  
 CC cell line. Codon optimisation results in improved expression of the  
 CC gagpol protein and reduces the risk of recombination between the transfer  
 CC vector and gagpol mRNA.  
 XX  
 SQ Sequence 1503 BP; 339 A; 530 C; 433 G; 201 T; 0 other;  
 Query Match 89.3%; Score 53.6; DB 21; Length 1503;  
 Best Local Similarity 93.3%; Pred. No. 2.2e-08;  
 Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 qacatcaagcagagcccccagagcccttcgcgactactgagccgtttcttcaagacc 60  
 Db 850 qacatcaagcagagcccccagagcccttcgcgactactgagccgtttcttcaagacc 909  
 RESULT 10  
 AAA70412  
 ID AAA70412 standard; DNA: 1515 BP.  
 XX  
 AC AAA70412;  
 DT 28-NOV-2000 (first entry)  
 DE Synthetic HIV Gag expression cassette coding sequence Gag.ModSP2.  
 XX  
 KW HIV-1; AIDS; Gag; vaccine; expression cassette; ss.  
 XX  
 OS Human immunodeficiency virus type 1.  
 OS Synthetic.  
 XX  
 PN W0200039302-A2.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 30-DEC-1999; 99WO-US31245.  
 PR 31-DEC-1998; 98US-0114495.  
 PR 01-DEC-1999; 99US-0168471.  
 XX  
 PA (CHIR.) CHIRON CORP.  
 XX  
 PI Barnett S, Zur Mequede J, Srivastava I, Lian Y, Bartoa K, Lin H;  
 PI Greer C, Selby M, Walker C;  
 XX  
 DR WP1; 2000-452400/39.  
 XX  
 PT Expression cassettes encoding the human immunodeficiency virus (HIV)  
 PT Gag containing polypeptide useful for vaccinating against HIV  
 PT infections and acquired immunodeficiency syndrome (AIDS) -  
 XX  
 PS claim 3; Fig 7; 39lpp; English.  
 XX  
 CC The present sequence is the coding sequence of a HIV Gag expression  
 CC cassette, Gag.ModSP2. The gag protein of HIV is needed for the assembly  
 CC of virus-like particles. In addition, the gag protein is involved in  
 CC many stages of the HIV life cycle, including assembly, virion maturation,  
 CC after particle release and early post-entry steps in viral replication.  
 CC The expression cassette may be used for the recombinant expression of  
 CC HIV Gag-polypeptides which may then be used to vaccinate against HIV  
 CC infection and acquired immunodeficiency syndrome (AIDS).  
 XX  
 SQ Sequence 1515 BP; 429 A; 547 C; 480 G; 159 T; 0 other;  
 Query Match 89.3%; Score 53.6; DB 21; Length 1515;  
 Best Local Similarity 93.3%; Pred. No. 2.2e-08;  
 Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 qacatcaagcagagcccccagagcccttcgcgactactgagccgtttcttcaagacc 60





Matches	56;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	1	gacataaagcaaggcccccaaggagcccttccggagactacgtgagaccgtttcttcaagacc	60						
Db	862	gacatacggccaggcccccaaggagcccttccggagactacgtgagaccgtttcttcaagacc	921						

Search completed. August 8, 2001, 18:24:58  
Job time: 12602 sec

GenCore version 4.5  
Copyright (c) 1993-2000 CompuGen Ltd.

em nucleic nucleic search, using sw model

Run on: August 8, 2001, 14:54:56 ; Search time 5719.75 Seconds  
(without alignments)  
162.256 Million cell updates/sec

Hit file: US-09-475-704-1

Perfect score: 60

Sequence: 1 qacatcaagcagggcccccda.....tggaccqcttcttcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 144157 seqs, 773874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenBank:

1: qb\_bal:\*

2: qb\_bal2:\*

3: qb\_bal3:\*

4: qb\_in1:\*

5: qb\_in2:\*

6: qb\_in3:\*

7: qb\_om:\*

8: qb\_ov:\*

9: qb\_pat1:\*

10: qb\_pat2:\*

11: qb\_ph:\*

12: qb\_pl1:\*

13: qb\_pl2:\*

14: qb\_pl3:\*

15: qb\_pl4:\*

16: em\_bal:\*

17: em\_bal2:\*

18: em\_fur1:\*

19: em\_htqo\_hum:\*

20: em\_htqo\_inv:\*

21: em\_htqo\_rnd:\*

22: em\_htq\_hum1:\*

23: em\_htq\_hum2:\*

24: em\_htq\_hum3:\*

25: em\_htq\_hum4:\*

26: em\_htq\_hum5:\*

27: em\_htq\_hum6:\*

28: em\_htq\_hum7:\*

29: em\_htq\_hum8:\*

30: em\_htq\_inv1:\*

31: em\_htq\_inv2:\*

32: em\_htq\_or\_her:\*

33: em\_htq\_rnd:\*

34: em\_hum1:\*

35: em\_hum2:\*

36: em\_hum3:\*

37: em\_hum4:\*

38: em\_hum5:\*

39: em\_hum6:\*

40: em\_hum7:\*

41: em\_in:\*

42: em\_om:\*

43: em\_or:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_pl:\*

48: em\_ro:\*

49: em\_sts:\*

50: em\_sy:\*

51: em\_un:\*

52: em\_v1:\*

53: qb\_sts1:\*

54: qb\_sts2:\*

55: qb\_sts3:\*

56: qb\_sy:\*

57: qb\_un:\*

58: qb\_v1:\*

59: qb\_v2:\*

60: qb\_htq1:\*

61: qb\_htq2:\*

62: qb\_htq3:\*

63: qb\_htq4:\*

64: qb\_htq5:\*

65: qb\_htq6:\*

66: qb\_htq7:\*

67: qb\_htq8:\*

68: qb\_htq9:\*

69: qb\_htq10:\*

70: qb\_htq11:\*

71: qb\_htq12:\*

72: qb\_htq13:\*

73: qb\_htq14:\*

74: qb\_htq15:\*

75: qb\_htq16:\*

76: qb\_htq17:\*

77: qb\_htq18:\*

78: qb\_htq19:\*

79: qb\_htq20:\*

80: qb\_htq21:\*

81: qb\_htq22:\*

82: qb\_htq23:\*

83: qb\_htq24:\*

84: qb\_htq25:\*

85: qb\_pr1:\*

86: qb\_pr2:\*

87: qb\_pr3:\*

88: qb\_pr4:\*

89: qb\_pr5:\*

90: qb\_pr6:\*

91: qb\_pr7:\*

92: qb\_pr8:\*

93: qb\_pr9:\*

94: qb\_ro1:\*

95: qb\_ro2:\*

96: qb\_in4:\*

97: qb\_pr10:\*

98: em\_bal3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	53.6	89.3	1509	56	AF201927	AF201927 Synthetic
2	53.6	89.3	1847	56	AF202464	AF202464 Synthetic
3	53.6	89.3	1847	56	AF202465	AF202465 Synthetic
4	50.4	84.0	1548	56	AF287354	AF287354 Synthetic
5	50.4	84.0	4352	56	AF287352	AF287352 Synthetic
6	50.4	84.0	4353	56	AF287353	AF287353 Synthetic
7	43	71.7	4307	9	AX019132	AX019132 Sequence
8	43	71.7	4307	9	AX035453	AX035453 Sequence











```
RESULT 12
AX035463
LOCUS      4642 bp      DNA
DEFINITION Sequence 12 from Patent WO0055341.
ACCESSION  AX035463
VERSION     AX035463.1  GI:11191105
KEYWORDS    synthetic construct.
SOURCE      synthetic construct.
ORGANISM    artificial construct.
REFERENCE   1 (bases 1 to 4642)
AUTHORS     Uden,M. and Mitrophanous,K.
TITLE       Anti-viral vectors
JOURNAL     Patent WO 0055341 A 12 21 SEP 2000;
UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) ; MIKROPHANOUS KIRILACUS
(US)
FEATURES
Source
1.4642
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="psYNGP2 - codon optimised HIV-1 gagpol with leader
sequence"
BASE COUNT 1219 a 1273 c 1389 g 762 t
ORIGIN

Query Match 71.7%; Score 43; DB 9; Length 4642;
Best Local Similarity 83.1%; Pred. No. 0.0065,
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 qacatcaagcaggcccccaggagcccttcgcgaactagtggaagcttcttcaagac 59
||||| 11 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1185 GACATCCGCCAAGCCGCCGAGGAAACCCCTTCGGGACTAGCTGGACCGCTTCTACAAAC 1243

RESULT 13
HIM274572
LOCUS      663 bp      DNA
DEFINITION Human immunodeficiency virus type 1 proviral partial gag gene for
P24, isolate 97SE-1189.
ACCESSION  AJ274572
VERSION     AJ274572.1  GI:7414161
KEYWORDS    gag gene; p24.
SOURCE      Human immunodeficiency virus type 1.
ORGANISM    human immunodeficiency virus type 1
REFERENCE   1 (bases 1 to 663)
AUTHORS     Toure-Kane,C., Montavon,C., Faye,M.A., Gueye,P.M., Sow,P.S.,
Ndoye,I., Gaye-Diallo,A., Delaporte,E., Peeters,M. and Mboup,S.
TITLE       Identification of all HIV type 1 group M subtypes in Senegal, a
country with low and stable seroprevalence
JOURNAL     AIDS Res. Hum. Retroviruses 16 (6), 603-609 (2000)
MEDLINE     20236909
REFERENCE   2 (bases 1 to 663)
AUTHORS     Peeters,M.
TITLE       Direct Submission
JOURNAL     Submitted (27 OCT-1999) Peeters M., Retrovirus, Ird, BP 5045, 34032
Montpellier cedex 1, FRANCE
FEATURES
Source
1..663
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="97SE-1189"
/db_xref="taxon:11676"
/country="Senegal"
1..663
/gene="gag"
<1..>663
/gene="gag"
/codon_start=1
gene
CDS

Query Match 63.7%; Score 38.2; DB 58; Length 393;
Best Local Similarity 78.0%; Pred. No. 0.28;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
/product="p24"
/protein_id="CAB85850.1"
/db_xref="GI:7414162"
/translation="QNIQGQMVYQATSPKTLNAWKVIEEKAFSPVLPMFSAIAEKA
TFQDINTMLNTVGGHQAAMQILKDTINEFAAEWDRLHPVHAGVPAGQITREPRGSLDIA
GTTSTLQEQITWMTSNPPIPVGEIYKGMIIILGNKIIVMYSPPVSLDIKQGPKEPRD
YVDRFFKTLRAEQSSQEVKSWMTDTLLIONANPDCKTILKALGSGASLEEMMTACQCV
GRP"
BASE COUNT 243 a 128 c 153 g 139 t
ORIGIN

Query Match 65.3%; Score 39.2; DB 59; Length 663;
Best Local Similarity 78.3%; Pred. No. 0.13;
Matches 47; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 qacatcaagcaggcccccaggagcccttcgcgaactagtggaagcttcttcaagac 60
||||| 11 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 445 GACATAAACAAAGGACAAAGGAAACCTTTAGAGACTATGTACACCGGTCTTCAAAACC 504

RESULT 14
AF106450
LOCUS      393 bp      DNA
DEFINITION HIV-1 isolate TZR5090 from Tanzania gag protein (gag) gene, partial
cds.
ACCESSION  AF106450
VERSION     AF106450.1  GI:6019003
KEYWORDS    Human immunodeficiency virus type 1.
SOURCE      Human immunodeficiency virus type 1
ORGANISM    Viruses; Retrovird viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE   1 (bases 1 to 393)
AUTHORS     Penjifo,P., Gilbert,P., Chaplin,B., Vannberg,F., Mwakagile,D.,
Msamanga,G., Hunter,D., Fawzi,W. and Essex,M.
TITLE       Emerging recombinant human immunodeficiency viruses: uneven
representation of the envelope V3 region
JOURNAL     AIDS 13 (13), 1613-1621 (1999)
MEDLINE     99437260
PUBMED     10509561
REFERENCE   2 (bases 1 to 393)
AUTHORS     Renjifo,B., Gilbert,P., Chaplin,B., Vannberg,F., Mwakagile,D.,
Fawzi,W. and Essex,M.
TITLE       Direct Submission
JOURNAL     Submitted (12-NOV-1998) Immunology and Infectious Diseases, Harvard
School of Public Health, 651 Huntington Avenue, Boston, MA 02115,
USA
FEATURES
Source
1..393
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="TZR5090"
/db_xref="taxon:11676"
/country="Tanzania"
<1..>393
/gene="gag"
<1..>393
/gene="gag"
/codon_start=1
/product="gag protein"
/protein_id="AAF01901.1"
/db_xref="GI:6019103"
/translation="KIVRMYSPPVSLIDIKQGPKEPRDYVDRFFKTLRAEQATQVKN
WMFFILLQVANPDCKNILLALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQANNA
NVMQRGNFKGPRRNKCFNCGEGHLAR"
BASE COUNT 145 a 70 c 97 g 81 t
ORIGIN

Query Match 63.7%; Score 38.2; DB 58; Length 393;
Best Local Similarity 78.0%; Pred. No. 0.28;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```









117: gb\_est48: \*  
 118: gb\_est49: \*  
 119: gb\_est50: \*  
 120: gb\_est51: \*  
 121: gb\_est52: \*  
 122: gb\_est53: \*  
 123: gb\_est54: \*  
 124: gb\_est55: \*  
 125: gb\_est56: \*  
 126: gb\_est57: \*  
 127: gb\_est58: \*  
 128: gb\_est59: \*  
 129: gb\_est60: \*  
 130: gb\_est61: \*  
 131: gb\_est62: \*  
 132: gb\_est63: \*  
 133: gb\_est64: \*  
 134: gb\_est65: \*  
 135: gb\_est66: \*  
 136: gb\_est67: \*  
 137: gb\_est68: \*  
 138: gb\_est69: \*  
 139: gb\_est70: \*  
 140: gb\_est71: \*  
 141: gb\_est72: \*  
 142: gb\_est73: \*  
 143: gb\_est74: \*  
 144: gb\_est75: \*  
 145: gb\_est76: \*  
 146: gb\_est77: \*  
 147: gb\_est78: \*  
 148: gb\_est79: \*  
 149: gb\_est80: \*  
 150: gb\_est81: \*  
 151: gb\_est82: \*  
 152: gb\_est83: \*  
 153: gb\_est84: \*  
 154: gb\_est85: \*  
 155: gb\_est86: \*  
 156: gb\_est87: \*  
 157: gb\_est88: \*  
 158: gb\_est89: \*  
 159: gb\_est90: \*  
 160: gb\_est91: \*  
 161: gb\_est92: \*  
 162: gb\_est93: \*  
 163: gb\_est94: \*  
 164: gb\_est95: \*  
 165: gb\_est96: \*  
 166: gb\_est97: \*  
 167: gb\_est98: \*  
 168: gb\_est99: \*  
 169: gb\_est100: \*  
 170: gb\_est101: \*  
 171: gb\_est102: \*  
 172: gb\_est103: \*  
 173: gb\_est104: \*  
 174: gb\_est105: \*  
 175: gb\_est106: \*  
 176: gb\_est107: \*  
 177: gb\_est108: \*  
 178: gb\_est109: \*  
 179: gb\_est110: \*  
 180: gb\_est111: \*  
 181: gb\_est112: \*  
 182: gb\_est113: \*  
 183: gb\_est114: \*  
 184: gb\_est115: \*  
 185: gb\_est116: \*  
 186: gb\_est117: \*  
 187: gb\_est118: \*  
 188: gb\_est119: \*  
 189: gb\_est120: \*

190: gb\_est110: \*  
 191: gb\_est111: \*  
 192: gb\_est112: \*  
 193: gb\_est113: \*  
 194: gb\_est114: \*  
 195: gb\_est115: \*  
 196: gb\_est116: \*  
 197: gb\_est117: \*  
 198: gb\_est118: \*  
 199: gb\_est119: \*  
 200: gb\_est120: \*  
 201: gb\_est121: \*  
 202: gb\_est122: \*  
 203: gb\_est123: \*  
 204: gb\_est124: \*  
 205: gb\_est125: \*  
 206: gb\_est126: \*  
 207: gb\_est127: \*  
 208: gb\_est128: \*  
 209: gb\_est129: \*  
 210: gb\_est130: \*  
 211: gb\_est131: \*  
 212: gb\_est132: \*  
 213: gb\_est133: \*  
 214: gb\_est134: \*  
 215: gb\_est135: \*  
 216: gb\_est136: \*  
 217: gb\_est137: \*  
 218: gb\_est138: \*  
 219: gb\_est139: \*  
 220: gb\_est140: \*  
 221: gb\_est141: \*  
 222: gb\_est142: \*  
 223: gb\_est143: \*  
 224: gb\_est144: \*  
 225: gb\_est145: \*  
 226: gb\_est146: \*  
 227: gb\_est147: \*  
 228: gb\_est148: \*  
 229: gb\_est149: \*  
 230: gb\_est150: \*  
 231: gb\_est151: \*  
 232: gb\_est152: \*  
 233: gb\_est153: \*  
 234: gb\_est154: \*  
 235: gb\_est155: \*  
 236: gb\_est156: \*  
 237: gb\_est157: \*  
 238: gb\_est158: \*  
 239: gb\_est159: \*  
 240: gb\_est160: \*  
 241: gb\_est161: \*  
 242: gb\_est162: \*  
 243: gb\_est163: \*  
 244: gb\_est164: \*  
 245: gb\_est165: \*  
 246: gb\_est166: \*  
 247: gb\_est167: \*  
 248: gb\_est168: \*  
 249: gb\_est169: \*  
 250: gb\_est170: \*  
 251: gb\_est171: \*  
 252: gb\_est172: \*  
 253: gb\_est173: \*  
 254: gb\_est174: \*  
 255: gb\_est175: \*  
 256: gb\_est176: \*  
 257: gb\_est177: \*  
 258: gb\_est178: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.









```

Seq primer: PolyTMix
High quality sequence start: 8
High quality sequence stop: 358
POLYA-No.

FEATURES
    source
        Location/Qualifiers
            1..364
                /organism="Sorghum bicolor"
                /db_xref="taxon:4558"
                /clone_lib "Ovary 1 (OVI)"
                /note="Organ: Mix of ovaries of varying immature stages
                        from 8-week-old plants; Vector: pBluescript II from Lambda
                        Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
                        from pCLY-A RNA in the cloning vector lambda Zap II.
                        Clones to be sequenced were prepared by mass excision."
            66 a      107 c      103 g      88 t

BASE COUNT
ORIGIN
    Query Match      46.7%;   Score 28;   DB 141;   Length 364;
    Best Local Similarity 66.7%;   Pred. No. 43;
    Matches 40;   Conservative 0;   Mismatches 20;   Indels 0;   Gaps 0;

```

QY	1	gacatcaaacgaagggcccaaaqqaagccttcgcgcgaactacgttqqacgctttcttcagaagacc	60
bB	75	GTCATCAAGGAGACGTGGCACCTAAACAAGTCTTGAGCATACTCCTCTTAAGAACC	144

RESULT 13

T14732/c

LOCUS T14732 364 bp mRNA EST 17 OCT 1996

DEFINITION O5c04107-121 membrane-free polysomes from endosperm Zea mays cDNA clone O5c04107 3' end, mRNA sequence.

ACCESSION T14732

VERSION T14732.1 GI:440711

KEYWORDS EST,

SOURCE zea mays.

ORGANISM zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

REFERENCE	1 (bases 1 to 364)
AUTHORS	Shen,B., Carneiro,N., Torres-Jerez,I., Stevenson,R., McCreery,T., Belentjaris,F., Baysdorfer,C., Almira,E., Perl,R., Hubben,J. and Larkins,B.
TITLE	Partial sequencing and mapping of clones from two maize cDNA libraries
JOURNAL	Plant Mol. Biol. 26, 1085-1101 (1994)
MEDLINE	95111093
COMMENT	Contact: The Maize cDNA Project  Belentjaris TG (primary contact) Dept. of Plant Sciences University of Arizona Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721 ph: 602-6218-746 fax: 602-621-7186 E-mail: belentjaris@ccit.arizona.edu

Chris Raysdorfer  
Department of Biological Sciences, School of Science  
California State University, Hayward  
Hayward, CA 94542  
ph: 510-881-4454  
fax: 510-727-2035  
E-mail: chaysdor@sl.csuhayward.edu

Rob Fertl  
Interdisciplinary Center for Biotechnology Research  
DNA Sequencing Core  
University of Florida  
P.O. Box 100695  
Gainesville, FL 32611-0695  
ph: 904-392-1928, ext. 301





[illegible]

Search completed: August 8, 2001, 20:58:57  
Job time: 21841 sec



EST: *	
1:	qb_est1: *
2:	qb_est2: *
3:	qb_est3: *
4:	qb_est4: *
5:	qb_est5: *
6:	qb_est6: *
7:	qb_est7: *
8:	qb_est8: *
9:	qb_est9: *
10:	qb_est10: *
11:	qb_est11: *
12:	qb_est12: *
13:	qb_est13: *
14:	qb_est14: *
15:	qb_est15: *
16:	qb_est16: *
17:	qb_est17: *
18:	qb_est18: *
19:	qb_est19: *
20:	qb_est20: *
21:	qb_est21: *
22:	qb_est22: *
23:	qb_est23: *
24:	qb_est24: *
25:	qb_est25: *
26:	qb_est26: *
27:	qb_est27: *
28:	qb_est28: *
29:	qb_est29: *
30:	qb_est30: *
31:	qb_est31: *
32:	qb_est32: *
33:	em_est33: *
34:	em_est34: *
35:	em_est35: *
36:	em_est36: *
37:	em_est37: *
38:	em_est38: *
39:	em_est39: *
40:	em_est40: *
41:	em_est41: *
42:	em_est42: *
43:	em_est43: *

117: qb\_est48: \*  
118: qb\_est49: \*  
119: qb\_est50: \*  
120: qb\_est51: \*  
121: qb\_est52: \*  
122: qb\_est53: \*  
123: qb\_est54: \*  
124: qb\_est55: \*  
125: qb\_est56: \*  
126: qb\_est57: \*  
127: qb\_est58: \*  
128: qb\_est59: \*  
129: qb\_est60: \*  
130: qb\_est61: \*  
131: qb\_est62: \*  
132: qb\_est63: \*  
133: qb\_est64: \*  
134: qb\_est65: \*  
135: qb\_est66: \*  
136: qb\_est67: \*  
137: qb\_est68: \*  
138: qb\_est69: \*  
139: qb\_est70: \*  
140: qb\_est71: \*  
141: qb\_est72: \*  
142: qb\_est73: \*  
143: qb\_est74: \*  
144: qb\_est75: \*  
145: qb\_est76: \*  
146: qb\_est77: \*  
147: qb\_est78: \*  
148: qb\_est79: \*  
149: qb\_est80: \*  
150: qb\_est81: \*  
151: qb\_est82: \*  
152: qb\_est83: \*  
153: qb\_est84: \*  
154: qb\_est85: \*  
155: qb\_est86: \*  
156: qb\_est87: \*  
157: qb\_est88: \*  
158: qb\_est89: \*  
159: qb\_est90: \*  
160: qb\_est91: \*  
161: qb\_est92: \*  
162: qb\_est93: \*  
163: qb\_est94: \*  
164: qb\_est95: \*  
165: qb\_est96: \*  
166: qb\_est97: \*  
167: qb\_est98: \*  
168: qb\_est99: \*  
169: qb\_est100: \*  
170: qb\_est101: \*  
171: qb\_est102: \*  
172: qb\_est103: \*  
173: qb\_est104: \*  
174: qb\_est105: \*  
175: qb\_est106: \*  
176: qb\_est107: \*  
177: qb\_est108: \*  
178: qb\_est109: \*  
179: qb\_est110: \*  
180: qb\_est111: \*  
181: qb\_est112: \*  
182: qb\_est113: \*  
183: qb\_est114: \*  
184: qb\_est115: \*  
185: qb\_est116: \*  
186: qb\_est117: \*  
187: qb\_est118: \*  
188: qb\_est119: \*  
189: qb\_est120: \*

190: qb\_est121: \*  
191: qb\_est122: \*  
192: qb\_est123: \*  
193: qb\_est124: \*  
194: qb\_est125: \*  
195: qb\_est126: \*  
196: qb\_est127: \*  
197: qb\_est128: \*  
198: qb\_est129: \*  
199: qb\_est130: \*  
200: qb\_est131: \*  
201: qb\_est132: \*  
202: qb\_est133: \*  
203: qb\_est134: \*  
204: qb\_est135: \*  
205: qb\_est136: \*  
206: qb\_est137: \*  
207: qb\_est138: \*  
208: qb\_est139: \*  
209: qb\_est140: \*  
210: qb\_est141: \*  
211: qb\_est142: \*  
212: qb\_est143: \*  
213: qb\_est144: \*  
214: qb\_est145: \*  
215: qb\_est146: \*  
216: qb\_est147: \*  
217: qb\_est148: \*  
218: qb\_est149: \*  
219: qb\_est150: \*  
220: qb\_est151: \*  
221: qb\_est152: \*  
222: qb\_est153: \*  
223: qb\_est154: \*  
224: qb\_est155: \*  
225: qb\_est156: \*  
226: qb\_est157: \*  
227: qb\_est158: \*  
228: qb\_est159: \*  
229: qb\_est160: \*  
230: qb\_est161: \*  
231: qb\_est162: \*  
232: qb\_est163: \*  
233: qb\_est164: \*  
234: qb\_est165: \*  
235: qb\_est166: \*  
236: qb\_est167: \*  
237: qb\_est168: \*  
238: qb\_est169: \*  
239: qb\_est170: \*  
240: qb\_est171: \*  
241: qb\_est172: \*  
242: qb\_est173: \*  
243: qb\_est174: \*  
244: qb\_est175: \*  
245: qb\_est176: \*  
246: qb\_est177: \*  
247: qb\_est178: \*  
248: qb\_est179: \*  
249: qb\_est180: \*  
250: qb\_est181: \*  
251: qb\_est182: \*  
252: qb\_est183: \*  
253: qb\_est184: \*  
254: qb\_est185: \*  
255: qb\_est186: \*  
256: qb\_est187: \*  
257: qb\_est188: \*  
258: qb\_est189: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.









```

/clone_lib="Pathogen induced 1 (PIL)"
/note="Organ: Anthracnose-infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
vector: pBluescript II from Lambda Zap II; Site_1: XhoI,
Site_2: EcoRI; Two-week-old sorghum plants (BTX 623
cultivar) were infected with pathogen (isolate FRM421 of
Colletotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings (2 weeks old) after inoculation. Note: young
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda ZAP II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen."
BASE COUNT      85 a      96 g      69 t
ORIGIN

Query Match      45.7%; Score 27.4; DB 137; Length 370;
Best Local Similarity 69.8%; Pred. No. 87;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 8 gccagggcccaaggagccctccgcactacgtggacgcttcttcagagcc 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 51 gccagggcacaatcagagccctccgcactacgtggacgcttcttcagagcc 103
    | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
BG462293      372 bp      mRNA      EST      20-MAY-2001
LOCUS      947046F01.y1 947 - 2 week shoot from Barkan lab 2ea mays cDNA, mRNA
DEFINITION      sequence.
ACCESSION      BG462293
VERSION      BG462293 ; GI:13388362
KEYWORDS      EST.
SOURCE      Zea mays.
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 372)
AUTHORS      Walbot,V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL      Unpublished (1999)
COMMENT      Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 947046 row: F column: 01.
FEATURES
    source
        location/Qualifiers
            1..372
            /organism="Zea mays"
            /cultivar="B73"
            /db_xref="taxon:4577"
            /clone_lib="947 - 2 week shoot from Barkan lab"
            /tissue_type="leaf and stem, including leaf base"
            /dev_stage="2 week old seedling (3 leaves)"
            /lab_host="XL1-Blue"
            /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);
            Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
            Stratagene's Unizap XR cDNA cloning kit with the 5' end
            at the EcoRI site. The library represents 8 x 10e5
            independent recombinant phage. The plants were greenhouse
            grown."
BASE COUNT      66 a      111 c      133 g      62 t

```

```

ORIGIN

Query Match      45.7%; Score 27.4; DB 154; Length 372;
Best Local Similarity 69.8%; Pred. No. 87;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 8 gccagggcccaaggagccctccgcactacgtggacgcttcttcagagcc 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 272 gccagggcacaatcagagccctccgcactacgtggacgcttcttcagagcc 324
    | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
BE364444      398 bp      mRNA      EST      20-JUL-2000
LOCUS      PIL_14_C08.bl_A002 Pathogen induced 1 (PIL) Sorghum bicolor cDNA,
DEFINITION      mRNA sequence.
ACCESSION      BE364444
VERSION      BE364444.1 GI:9306001
KEYWORDS      EST.
SOURCE      Sorghum.
ORGANISM      Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE      1 (bases 1 to 398)
AUTHORS      Cordonnier-Pratt,M.M., Gingle,A., Dean,R., Sudman,M. and Pratt
,L.H.
TITLE      An EST database from Sorghum pathogen-induced plants
JOURNAL      Unpublished (2000)
COMMENT      Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 396
POLYA-NO.
FEATURES
    source
        location/Qualifiers
            1..398
            /organism="Sorghum bicolor"
            /db_xref="taxon:4558"
            /clone_lib="Pathogen induced 1 (PIL)"
            /note="Organ: Anthracnose-infected leaves from
            two-week-old sorghum plants 48 hr after inoculation;
            Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;
            Site_2: EcoRI; Two-week-old sorghum plants (BTX 623
            cultivar) were infected with pathogen (isolate FRM421 of
            Colletotrichum graminicola, which is a sorghum isolate).
            RNA was prepared from infected leaves harvested from 45
            seedlings 48 hours after inoculation. Note: young
            seedlings (2 weeks old) exhibit juvenile resistant
            reaction, which is an incompatible interaction. As they
            grow older (4 weeks or older), plants resume susceptibility
            to anthracnose disease. The library was made from poly-A
            RNA in the cloning vector lambda ZAP II. Clones to be
            sequenced were prepared by mass excision. WARNING: While
            most or all ESTs are expected to derive from the host
            plant, no effort was made to eliminate ESTs deriving from
            the pathogen."
BASE COUNT      68 a      128 c      136 g      66 t
ORIGIN

Query Match      45.7%; Score 27.4; DB 166; Length 398;
Best Local Similarity 69.8%; Pred. No. 88;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```



```
/dev_stage="2 week old seedling (3 leaves)"
/lab_host="XLI-Blue"
/note="Organ: shoot, Vector: Lambda ZAP (pBluescript SK-);
Site_1: EcoRI, Site_2: XhoI. Directionally cloned using
Stratagene's Unizap XR cDNA cloning kit with the 5' end
at the EcoRI site. The library represents 8 x 10e5
independent recombinant phage. The plants were greenhouse
grown."

BASE COUNT      87 a  156 c  136 g   76 t
ORIGIN

Query Match:
Best local Similarity  45.7%; Score 27.4; DB 155; Length 455;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY  8 gccagggccccagggcccttcgggactagtggaagcgtttctcaagacc 60
    1 1111 1 1111 1 111 1 11 1111 1 11111111 1
Db 174 GGCAGGACATCAAGAGACCCCGGAGGTCACAGCAGCCCATCTTCAAGAGC 226

RESULT 13
LOCUS BE997330/c
DEFINITION BE997330 470 bp mRNA EST 06-OCT-2000
          947017H06.x3 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
          sequence.
ACCESSION BE997330
VERSION BE997330 1 GI:10697606
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 470)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
        Contact: Walbot V
        Department of Biological Sciences
        Stanford University
        855 California Ave, Palo Alto, CA 94304, USA
        Tel: 650 723 2227
        Fax: 650 725 8221
        Email: walbot@stanford.edu
        Plate: 947017 row: H column: 06.
        Location/Qualifiers
          1..470
            /organism="Zea mays"
            /cultivar="B73"
            /db_xref="taxon:4577"
            /clone_lib="947 - 2 week shoot from Barkan lab"
            /tissue_type="leaf and stem, including leaf base"
            /dev_stage="2 week old seedling (3 leaves)"
            /lab_host="XLI-Blue"
            /note="Organ: shoot, Vector: Lambda ZAP (pBluescript SK-);
            Site_1: EcoRI, Site_2: XhoI. Directionally cloned using
            Stratagene's Unizap XR cDNA cloning kit with the 5' end
            at the EcoRI site. The library represents 8 x 10e5
            independent recombinant phage. The plants were greenhouse
            grown."

BASE COUNT      90 a  135 c  144 g  101 t
ORIGIN

Query Match:
Best local Similarity  45.7%; Score 27.4; DB 143; Length 470;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY  8 gccagggccccagggcccttcgggactagtggaagcgtttctcaagacc 60
    1 1111 1 1111 1 111 1 11 1111 1 11111111 1
Db 422 GGCAGGACATCAAGAGACCCCGGAGGTCACAGCAGCCCATCTTCAAGAGC 370

/dev_stage="2 week old seedling (3 leaves)"
/lab_host="XLI-Blue"
/note="Organ: shoot, Vector: Lambda ZAP (pBluescript SK-);
Site_1: EcoRI, Site_2: XhoI. Directionally cloned using
Stratagene's Unizap XR cDNA cloning kit with the 5' end
at the EcoRI site. The library represents 8 x 10e5
independent recombinant phage. The plants were greenhouse
grown."

BASE COUNT      89 a  135 c  138 g  108 t
ORIGIN

Query Match:
Best local Similarity  45.7%; Score 27.4; DB 152; Length 470;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY  8 gccagggccccagggcccttcgggactagtggaagcgtttctcaagacc 60
    1 1111 1 1111 1 111 1 11 1111 1 11111111 1
Db 457 GGCAGGACATCAAGAGACCCCGGAGGTCACAGCAGCCCATCTTCAAGAGC 405

RESULT 15
LOCUS AI973498/c
DEFINITION AI973498 478 bp mRNA EST 25-AUG-1999
          496027Bi2.x1 496 - stressed shoot cDNA library from Wang/Bohnert
          lab Zea mays cDNA, mRNA sequence.
ACCESSION AI973498
VERSION AI973498.1 GI:5770324
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 478)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
```







PS Claim 1: Page 92; 113pp; English  
XX Expression cassettes comprising a polynucleotide encoding antigenic  
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are  
CC useful in DNA immunization, generation of packaging cell lines and  
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag  
CC expression cassettes exhibit increased potency for induction of  
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1  
CC self-assemble into non-infectious virus-like particles which are used as  
CC a matrix for the proper presentation of an antigen entrapped or  
CC associated to the immune system of the host.  
XX  
SQ Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 60;  
Best Local Similarity 100.0%; Pred No 7 20-10;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatccgcagggccccaaggagcccttcgcgactactggtgacgctttcttcaagacc 60  
|||||  
Db 1 gacatccgcagggccccaaggagcccttcgcgactactggtgacgctttcttcaagacc 60

RESULT 2  
AAA51610  
ID AAA51610 standard; DNA; 1509 BP.  
XX  
AC AAA51610;  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE HIV synthetic Gag polynucleotide.  
XX  
KW Gag; expression cassette; antigenic, type C; HIV; Env; synthetic;  
KW DNA immunization; packaging cell line; antigen presentation; ss.  
XX  
XX Human immunodeficiency virus type C strain AF110967.  
OS  
OS Synthetic.  
XX  
PN WO200039304-A2.  
XX  
PD 06-JUL-2000.  
XX  
PF 30-DEC-1999; 99WO-US31273.  
XX  
PR 31-DEC-1998; 99US-0114495.  
PR 01-SEP-1999; 99US-0152195.  
XX  
PA (CHIR ) CHIRON CORP  
XX  
PI Barnett S, Zur Megede J;  
XX  
DP WPI: 2000-452401/39.  
XX  
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV  
PT Env polypeptide and the polypeptide useful for immunizing a mammal  
PT especially human against HIV  
XX

Claim 2: Page 93; 113pp; English  
XX Expression cassettes comprising a polynucleotide encoding antigenic  
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are  
CC useful in DNA immunization, generation of packaging cell lines and  
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag  
CC expression cassettes exhibit increased potency for induction of  
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1  
CC self-assemble into non-infectious virus-like particles which are used as  
CC a matrix for the proper presentation of an antigen entrapped or  
CC associated to the immune system of the host.  
XX  
SQ Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 1509;  
Best Local Similarity 100.0%; Pred No 9e-10;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gacatccgcagggccccaaggagcccttcgcgactactggtgacgctttcttcaagacc 60  
|||||  
Db 841 gacatccgcagggccccaaggagcccttcgcgactactggtgacgctttcttcaagacc 900

RESULT 3  
AAA51626  
ID AAA51626 standard; DNA; 1509 BP.  
XX  
AC AAA51626;  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE HIV codon-optimized synthetic Gag polynucleotide.  
XX  
KW Gag; expression cassette; antigenic, type C; HIV; Env; synthetic;  
KW DNA immunization; packaging cell line; antigen presentation; ss.  
XX  
XX Human immunodeficiency virus type C strain AF110967.  
OS  
OS Synthetic.  
XX  
PN WO200039304-A2.  
XX  
PD 06-JUL-2000  
XX  
PF 30-DEC-1999; 99WO-US31273.  
XX  
PR 31-DEC-1998; 98US-0114495.  
PR 01-SEP-1999; 99US-0152195.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Barnett S, Zur Megede J;  
XX  
DP WPI: 2000-452401/39.  
XX  
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV  
PT Env polypeptide and the polypeptide useful for immunizing a mammal  
PT especially human against HIV  
XX  
XX Disclosure; Page 104, 113pp; English.  
XX  
CC Expression cassettes comprising a polynucleotide encoding antigenic  
CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are  
CC useful in DNA immunization, generation of packaging cell lines and  
CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag  
CC expression cassettes exhibit increased potency for induction of  
CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1  
CC self-assemble into non-infectious virus-like particles which are used as  
CC a matrix for the proper presentation of an antigen entrapped or  
CC associated to the immune system of the host.  
XX  
SQ Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 1509;  
Best Local Similarity 100.0%; Pred No 9e-10;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gacatccgcagggccccaaggagcccttcgcgactactggtgacgctttcttcaagacc 60  
|||||  
Db 841 gacatccgcagggccccaaggagcccttcgcgactactggtgacgctttcttcaagacc 900

RESULT 4  
AAA70426  
ID AAA70426 standard; DNA; 60 BP.  
XX





XX 23-MAR-2000.  
PD  
XX 10-SEP-1999; 99WO-US20675.  
PF  
XX 11-SEP-1998; 98US-0100022. 60  
PR 12-SEP-1998; 98US-0100063. 60  
XX  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
XX Gray JT, Mulligan RC;  
XX WPI: 2000-271455/23.  
XX P-PSDB; AAY70599.  
XX  
XX New packaging cell line for producing a viral accessory protein  
PT independent HIV derived retroviral vector particles, useful in gene  
PT therapy or gene replacement -  
XX  
XX Claim 41; Fig 8; 62pp; English.  
XX  
XX The patent discloses new packaging cell line for producing a viral  
CC accessory protein independent lentivirus, preferably  
CC human immunodeficiency virus (HIV), derived retroviral  
CC The packaging cell line comprises a mammalian cell, a retroviral DNA  
CC comprising a coding sequence for a lentivirus, preferably HIV, gagpol,  
CC where the coding sequence has been mutagenised to improve expression of  
CC the viral gagpol proteins, a second retroviral nucleotide sequence  
CC comprising the coding sequence for a heterologous envelope protein and a  
CC third retroviral nucleotide sequence comprising a DNA sequence of  
CC interest and lentivirus, preferably HIV, cis-acting sequences required  
CC for packaging, reverse transcription and integration.  
CC The packaging cell lines and viral particles can be used for gene  
CC therapy or gene replacement with improved safety. They can also be used  
CC in the development and production of vaccines and biochemical reagents.  
CC The present sequence is a gag coding region of codon optimised HIV  
CC gagpol sequence. This sequence is used in the packaging  
CC cell line. Codon optimisation results in improved expression of the  
CC gagpol protein and reduces the risk of recombination between the transfer  
CC vector and gagpol mRNA.  
XX  
XX Sequence 1503 BP; 339 A, 530 C, 433 G; 201 T; 0 other;  
SQ

Query Match 97.3%; Score 58.4; DB 21; Length 1503;  
Best Local Similarity 98.3%; P-Id No. 2.7e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gacatccgcagggcccccagggcccttcgcgactactggtgacccgttcttcaagacc 60  
|||||  
Db 850 gacatccgcagggcccccagggcccttcgcgactactggtgacccgttcttcaagacc 909

RESULT 7  
AAA70412  
ID AAA70412 standard; DNA; 1515 BP.  
XX  
XX AAA70412;  
AC  
XX 28-NOV-2000 (first entry)  
DT  
XX Synthetic HIV Gag expression cassette coding sequence Gag.ModSF2.  
DE  
XX HIV-1; AIDS; Gag; vaccine; expression cassette; ss.  
KW  
XX Human immunodeficiency virus type 1.  
OS  
XX Synthetic.  
XX  
XX WO200039302-A2.  
PN  
XX 06-JUL-2000  
PD  
XX  
XX AAA70412;  
AC  
XX 28-NOV-2000 (first entry)  
DT  
XX Synthetic HIV Gag expression cassette coding sequence Gag.ModSF2.  
DE  
XX HIV-1; AIDS; Gag; vaccine; expression cassette; ss.  
KW  
XX Human immunodeficiency virus type 1.  
OS  
XX Synthetic.  
XX  
XX WO200039302-A2.  
PN  
XX 06-JUL-2000  
PD  
XX  
XX 30-DEC-1999; 99WO-US31245.  
PF

XX 31-DEC-1998; 98US-0114495.  
PR 01-DEC-1999; 99US-0168471.  
XX  
XX (CHIR ) CHIRON CORP.  
XX Barnett S, Zur Meqede J, Srivastava I, Lian Y, Hartog K, Liu H;  
PI Greer C, Selby M, Walker C;  
XX WPI: 2000-452400/39.  
DR  
XX Expression cassettes encoding the human immunodeficiency virus (HIV)  
XX Gag-containing polypeptide useful for vaccinating against HIV  
PT infections and acquired immunodeficiency syndrome (AIDS) -  
PT  
XX Claim 3; Fig 7; 391pp, English.  
PS  
XX The present sequence is the coding sequence of a HIV Gag expression  
CC cassette, Gag.ModSF2. The Gag protein of HIV is needed for the assembly  
CC of virus-like particles. In addition, the Gag protein is involved in  
CC many stages of the HIV life cycle, including assembly, virion maturation  
CC after particle release and early post-entry steps in viral replication.  
CC The expression cassette may be used for the recombinant expression of  
CC HIV Gag-polypeptides which may then be used to vaccinate against HIV  
CC infection and acquired immunodeficiency syndrome (AIDS).  
XX  
XX Sequence 1515 BP; 429 A; 547 C; 480 G; 159 T; 0 other;  
SQ

Query Match 97.3%; Score 58.4; DB 21; Length 1515;  
Best Local Similarity 98.3%; P-Id No. 2.7e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gacatccgcagggcccccagggcccttcgcgactactggtgacccgttcttcaagacc 60  
|||||  
Db 862 gacatccgcagggcccccagggcccttcgcgactactggtgacccgttcttcaagacc 921

RESULT 8  
AAA70413  
ID AAA70413 standard; DNA; 1853 BP.  
XX  
XX AAA70413;  
AC  
XX 28-NOV-2000 (first entry)  
DT  
XX HIV Gag-polypease expression cassette coding sequence Gagprot Mods.  
DE  
XX HIV-1; AIDS; Gag protease; vaccine; expression cassette; ss.  
KW  
XX Human immunodeficiency virus type 1.  
OS  
XX Synthetic.  
XX  
XX WO200039302-A2.  
PN  
XX 06-JUL-2000.  
PD  
XX  
XX 30-DEC-1999; 99WO-US31245.  
PF  
XX  
XX 31-DEC-1998; 98US-0114495.  
PR 01-DEC-1999; 99US-0168471.  
XX  
XX (CHIR ) CHIRON CORP.  
XX Barnett S, Zur Meqede J, Srivastava I, Lian Y, Hartog K, Liu H;  
PI Greer C, Selby M, Walker C;  
XX WPI: 2000-452400/39.  
DR  
XX Expression cassettes encoding the human immunodeficiency virus (HIV)  
PT Gag-containing polypeptide useful for vaccinating against HIV  
PT infections and acquired immunodeficiency syndrome (AIDS) -  
PT  
XX



XX 28-NOV-2000 (first entry)  
DT Synthetic HIV Gag/HCV core fusion coding sequence.  
DE HIV-1; AIDS; Gag; vaccine; expression cassette, ss.  
XX Chimeric - Human immunodeficiency virus type 1.  
OS Chimeric - Hepatitis C virus.  
XX WO200039302-A2.  
PN 06-JUL-2000.  
PD 30-DEC-1999; 99WO-US31245.  
XX 31-DEC-1998; 98US-0114495.  
PR 01-DEC-1999; 99US-0168471.  
XX (CHIR ) CHIRON CORP.  
XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H,  
PI Greer C, Selby M, Walker C;  
XX WPI; 2000-452400/39.  
XX Expression cassettes encoding the human immunodeficiency virus (HIV)  
Gag-containing polypeptide useful for vaccinating against HIV  
infectious and acquired immunodeficiency syndrome (AIDS) -  
XX Example 1; Pages 341-342; 391pp; English.  
XX The present sequence is a HIV Gag/Hepatitis C virus (HCV) core fusion  
coding sequence. The Gag protein of HIV is needed for the assembly of  
virus-like particles. In addition, the Gag protein is involved in many  
stages of the HIV life cycle, including assembly, virion maturation after  
particle release and early post-entry steps in viral replication. The  
present invention relates to synthetic HIV Gag expression cassettes. The  
present sequence was cloned and used to generate the expression cassettes  
of the present invention. The expression cassettes may be used for the  
recombinant expression of HIV Gag polypeptides which may then be used to  
vaccinate against HIV infection and acquired immunodeficiency syndrome  
(AIDS).  
XX Sequence 2031 BP; 421 A; 707 C; 645 G; 257 T; 0 other;  
SQ  
Query Match 97.3%; Score 58.4; DB 21; Length 2031;  
Best Local Similarity 98.3%; Pred. No. 2.8e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gacatccgcagggcccaaggagcccttcgcgactactgagccgctttctacaagacc 60  
|||||  
Db 862 gacatccgcagggcccaaggagcccttcgcgactactgagccgctttctacaagacc 921  
RESULT 12  
AAA70414  
ID AAA70414 standard; DNA; 4319 BP.  
XX  
AC AAA70414;  
XX 28-NOV-2000 (first entry)  
DT HIV Gag-polymerase expression cassette coding sequence GagPol.ModSF.  
DE HIV-1; AIDS; Gag-polymerase; vaccine; expression cassette; ss.  
XX Human immunodeficiency virus type 1.  
OS Synthetic.  
XX WO200039302-A2.  
PN

PD 06-JUL-2000.  
XX 30-DEC-1999; 99WO-US31245.  
XX 31-DEC-1998; 98US-0114495.  
PR 01-DEC-1999; 99US-0168471.  
XX (CHIR ) CHIRON CORP.  
XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H,  
PI Greer C, Selby M, Walker C;  
XX WPI; 2000-452400/39.  
XX Expression cassettes encoding the human immunodeficiency virus (HIV)  
Gag-containing polypeptide useful for vaccinating against HIV  
infectious and acquired immunodeficiency syndrome (AIDS) -  
XX Claim 10; Fig 7; 391pp; English.  
XX The present sequence is the coding sequence of a HIV Gag-polymerase  
expression cassette, GagPol.ModSF. The Gag protein of HIV is needed for  
the assembly of virus-like particles. In addition, the Gag protein is  
involved in many stages of the HIV life cycle, including assembly, virion  
maturation after particle release and early post entry steps in viral  
replication. The expression cassette may be used for the recombinant  
expression of HIV Gag polypeptides which may then be used to vaccinate  
against HIV infection and acquired immunodeficiency syndrome (AIDS).  
XX Sequence 4319 BP; 1003 A; 1426 C; 1361 G; 529 T; 0 other;  
SQ  
Query Match 97.3%; Score 58.4; DB 21; Length 4319;  
Best Local Similarity 98.3%; Pred. No. 3e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gacatccgcagggcccaaggagcccttcgcgactactgagccgctttctacaagacc 60  
|||||  
Db 862 gacatccgcagggcccaaggagcccttcgcgactactgagccgctttctacaagacc 921  
RESULT 13  
AAA70472  
ID AAA70472 standard; DNA; 4472 BP.  
XX  
AC AAA70472;  
XX 28-NOV-2000 (first entry)  
DT HIV bicistronic construct gfp100 modUS4 delV1/V2 Gag.ModSF2.  
DE HIV-1; AIDS; Gag; vaccine; expression cassette; Env; ss.  
XX Chimeric - Cytomegalovirus  
OS Chimeric - Human immunodeficiency virus type 1.  
XX WO200039302-A2.  
PN 06-JUL-2000.  
XX 30-DEC-1999; 99WO-US31245.  
XX 31-DEC-1998; 98US-0114495.  
PR 01-DEC-1999; 99US-0168471.  
XX (CHIR ) CHIRON CORP.  
XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H,  
PI Greer C, Selby M, Walker C;  
XX WPI; 2000-452400/39.  
XX Expression cassettes encoding the human immunodeficiency virus (HIV)



Query Match 97.3%; Score 58.4; DB 21; Length 4689;  
Best Local Similarity 98.3%; Pred. No. 3e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 gacatccgcacaggcccccaaggagcccttcgggactactggtggaacgctttcttcaagacc 60  
|||||  
Db 4011 gacatccgcacaggcccccaaggagcccttcgggactactggtggaacgctttcttcaagacc 4070  
|||||

Search completed: August 8, 2001, 18:25:01  
Job time: 12605 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

em nucleic - nucleic search, using sw model

Run on: August 8, 2001, 20:00:32 ; Search time 5719.75 Seconds  
(without alignments)  
162.256 Million cell updates/sec

File: us-09-475-704-2  
Perfect score: 60  
Sequence: 1 gacatccgacagggcccccacaa.....tggacccgcttcttcaaac 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 1444157 seqs, 773874588 residues 2688314

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank1:  
1: qb\_bal: \*  
2: qb\_bal2: \*  
3: qb\_bal3: \*  
4: qb\_inl: \*  
5: qb\_in2: \*  
6: qb\_in3: \*  
7: qb\_in4: \*  
8: qb\_ov: \*  
9: qb\_pat1: \*  
10: qb\_pat2: \*  
11: qb\_ph: \*  
12: qb\_pl1: \*  
13: qb\_pl2: \*  
14: qb\_pl3: \*  
15: qb\_pl4: \*  
16: em\_bal: \*  
17: em\_bal2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_rod: \*  
22: em\_htq\_hum1: \*  
23: em\_htq\_hum2: \*  
24: em\_htq\_hum3: \*  
25: em\_htq\_hum4: \*  
26: em\_htq\_hum5: \*  
27: em\_htq\_hum6: \*  
28: em\_htq\_hum7: \*  
29: em\_htq\_hum8: \*  
30: em\_htq\_inv1: \*  
31: em\_htq\_inv2: \*  
32: em\_htq\_other: \*  
33: em\_htq\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_em: \*  
43: em\_or: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_ro: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_vl: \*  
53: qb\_sts1: \*  
54: qb\_sts2: \*  
55: qb\_sts3: \*  
56: qb\_sy: \*  
57: qb\_un: \*  
58: qb\_vil: \*  
59: qb\_vl2: \*  
60: qb\_htq1: \*  
61: qb\_htq2: \*  
62: qb\_htq3: \*  
63: qb\_htq4: \*  
64: qb\_htq5: \*  
65: qb\_htq6: \*  
66: qb\_htq7: \*  
67: qb\_htq8: \*  
68: qb\_htq9: \*  
69: qb\_htq10: \*  
70: qb\_htq11: \*  
71: qb\_htq12: \*  
72: qb\_htq13: \*  
73: qb\_htq14: \*  
74: qb\_htq15: \*  
75: qb\_htq16: \*  
76: qb\_htq17: \*  
77: qb\_htq18: \*  
78: qb\_htq19: \*  
79: qb\_htq20: \*  
80: qb\_htq21: \*  
81: qb\_htq22: \*  
82: qb\_htq23: \*  
83: qb\_htq24: \*  
84: qb\_htq25: \*  
85: qb\_pr1: \*  
86: qb\_pr2: \*  
87: qb\_pr3: \*  
88: qb\_pr4: \*  
89: qb\_pr5: \*  
90: qb\_pr6: \*  
91: qb\_pr7: \*  
92: qb\_pr8: \*  
93: qb\_pr9: \*  
94: qb\_rol: \*  
95: qb\_rod2: \*  
96: qb\_in4: \*  
97: qb\_pr10: \*  
98: em\_bal3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	58.4	97.3	1509	56	AF201927 Synthetic
2	58.4	97.3	1917	56	AF202464 Synthetic
3	58.4	97.3	1847	56	AF202465 Synthetic
4	48.8	81.3	1548	56	AF287354 Synthetic
5	48.8	81.3	4352	56	AF287352 Synthetic
6	48.8	81.3	4353	56	AF287353 Synthetic
7	47.8	79.7	4307	9	AX019132 Sequence
8	47.8	79.7	4307	9	AX035453 Sequence







```
/codon_start=1
/transl_table=11
/product="gag protein"
/protein_id="AAG28735.1"
/db_xref="GI:11066862"
/translation="MGARASVLSAGELDKWEKILRLPGGKKQYRLKHIVWASRELEFR
AVDPGLLETSEGCQILGLQIQLPSLQTSSEELSLNTVATLYCVHQKIEVKDTKEALE
KIEEONKSKKKAQAAADTGNSSQVSONYPIVONLQGMVHQAISPERTLNWVKVVE
EKAFSEVIMFSAISSEGAIPDNLNMLNIVGGHQAAMQMLKETINEEAAEWDRHPV
HAGPIAPGMREPRGSDIAGTTSTLQFOIGWMTNPPIPVGEIYKRWIILGLNKIVRM
YSPSTLIDIRQGPKEPRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCKTIL
KALGPAATLEEMMTACQGVGPGHKARVLAEMSOVNTSATIMMORGNFRNQRKTKC
FNCGKEGHIANKAPPKKGWKQKGEHQMKIATIEKANFLREDLAFPGKARKFSS
EQTRANSPIRERQVWRDNNLSSEAGADKQIVSPSPFQIILWDRPIVITIKIGGQIK
EALDGTADITVLEEMNPGMKMIGGIGGFIVKQYDQIPIEIGCHKAIGTVLVG
PTPVNIIGRNLLTOIGCTLNFPISPIETVPVVKLPGKMDGPKVQKWPPIEEKIKALVEI
CTEMEKEGKISKIGPENYNTPVFAIKKDKSTKWKRLVDRELNKRRTQDFWEVQLGIP
HPAGLKKKSVTLVDGDAYFSVPLHEDFKYTAFTIPSTINNETPGTYQYNNIPVGGW
KQSPAFQSSMTTILEPFRKQNDLVLYQYMDIIVYGSLELGHRTKIEELRQHLIR
WGFTTPDKHOKPEPPFLWGYELHPDKWTVQIVLPEKDSWIVNDQKLVGLNWSO
IYAGIKVPQLKLI PGTKALTEV IPLTEAELELAENKELKEPVHGVYVYDPSKDLIA
ETQKQCOQWYQIYQEPFNKLTGYAPRTGATNDVKQLEAVQKIATESIVINCK
TPKPLPIQKETWETWTEWQATWIPWFVNTPPVVKIWLQLEKEPIIGAEFEVD
GAANPETKSKAGVYTNKQKQVSLDTNQLTQLEQALYLAQDSGLFVNIVDSQY
ALGIIQAOPRSESELSVQIEQLIKKEKYLAWVPAHKGIGNEQVVKLVSAIRKV
LFLDGLQYACQEPFKYHNSNPAMASDFNI PPVAKFIVASQDKQIKGEAMHGVDCS
PGIWLQDCTHLEKVLVAVHVASGYIEAEVIPAETGQETAYFLKLAGRWPVTTHI
DNGSNFTSATVKAACWAGIKQEGIPYNPQSQGVVSEMNKELKLIQVVKQAEHLK
TAVQMAVFIHNEFKRGGIGGYSAGERIVDIATDIQKLELOKQITKQNFVYVYRDSR
DPLWKGPAKLLKQFCGAVVYQNSDIKVVPRRKAIIIRDYKQMGAGDDCVAGQDED"
BASE COUNT 1051 a 1358 c 1435 g 509 t
ORIGIN
```

```
Query Match 81.3%; Score 48.8; DB 56; Length 4353;
Best Local Similarity 88.3%; Pred No 0.00023;
Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatccgagggcccccaagagagcccttcgcgagactacatcaggagcccttccttcagagac 60
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
DB 886 GACATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 945

RESULT 6
AF287353 4353 bp DNA SYN 01-NOV-2000
LOCUS Synthetic construct HIV-1-derived gag-pol fusion protein gene,
DEFINITION complete cds.
ACCESSION AF287353
VERSION AF287353.1 GI:11066864
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Fuller, M. and Anson, D.S.
TITLE Helper plasmids for production of HIV derived vectors
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4353)
AUTHORS Fuller, M. and Anson, D.S.
TITLE Direct Submission
JOURNAL Submitted (12-Nov-2000) GenBank, North Adelaide, SA 5006, Australia
FEATURES
Location/Qualifiers
```

```
source
1. 4353
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="derived from Human immunodeficiency virus type 1
isolate YU-2"
37. 4344
/note="codon optimized reading frame"
/transl_table=11
/product="gag-pol fusion protein"
/protein_id="AAG28737.1"
/db_xref="GI:11066865"
/translation="MGARASVLSAGELDKWEKILRLPGGKKQYRLKHIVWASRELEFR
AVDPGLLETSEGCQILGLQIQLPSLQTSSEELSLNTVATLYCVHQKIEVKDTKEALE
KIEEONKSKKKAQAAADTGNSSQVSONYPIVONLQGMVHQAISPERTLNWVKVVE
EKAFSEVIMFSAISSEGAIPDNLNMLNIVGGHQAAMQMLKETINEEAAEWDRHPV
HAGPIAPGMREPRGSDIAGTTSTLQFOIGWMTNPPIPVGEIYKRWIILGLNKIVRM
YSPSTLIDIRQGPKEPRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCKTIL
KALGPAATLEEMMTACQGVGPGHKARVLAEMSOVNTSATIMMORGNFRNQRKTKC
FNCGKEGHIANKAPPKKGWKQKGEHQMKIATIEKANFLREDLAFPGKARKFSS
EQTRANSPIRERQVWRDNNLSSEAGADKQIVSPSPFQIILWDRPIVITIKIGGQIK
EALDGTADITVLEEMNPGMKMIGGIGGFIVKQYDQIPIEIGCHKAIGTVLVG
PTPVNIIGRNLLTOIGCTLNFPISPIETVPVVKLPGKMDGPKVQKWPPIEEKIKALVEI
CTEMEKEGKISKIGPENYNTPVFAIKKDKSTKWKRLVDRELNKRRTQDFWEVQLGIP
HPAGLKKKSVTLVDGDAYFSVPLHEDFKYTAFTIPSTINNETPGTYQYNNIPVGGW
KQSPAFQSSMTTILEPFRKQNDLVLYQYMDIIVYGSLELGHRTKIEELRQHLIR
WGFTTPDKHOKPEPPFLWGYELHPDKWTVQIVLPEKDSWIVNDQKLVGLNWSO
IYAGIKVPQLKLI PGTKALTEV IPLTEAELELAENKELKEPVHGVYVYDPSKDLIA
ETQKQCOQWYQIYQEPFNKLTGYAPRTGATNDVKQLEAVQKIATESIVINCK
TPKPLPIQKETWETWTEWQATWIPWFVNTPPVVKIWLQLEKEPIIGAEFEVD
GAANPETKSKAGVYTNKQKQVSLDTNQLTQLEQALYLAQDSGLFVNIVDSQY
ALGIIQAOPRSESELSVQIEQLIKKEKYLAWVPAHKGIGNEQVVKLVSAIRKV
LFLDGLQYACQEPFKYHNSNPAMASDFNI PPVAKFIVASQDKQIKGEAMHGVDCS
PGIWLQDCTHLEKVLVAVHVASGYIEAEVIPAETGQETAYFLKLAGRWPVTTHI
DNGSNFTSATVKAACWAGIKQEGIPYNPQSQGVVSEMNKELKLIQVVKQAEHLK
TAVQMAVFIHNEFKRGGIGGYSAGERIVDIATDIQKLELOKQITKQNFVYVYRDSR
DPLWKGPAKLLKQFCGAVVYQNSDIKVVPRRKAIIIRDYKQMGAGDDCVAGQDED"
BASE COUNT 1051 a 1358 c 1435 g 509 t
ORIGIN

Query Match 81.3%; Score 48.8; DB 56; Length 4353;
Best Local Similarity 88.3%; Pred No 0.00023;
Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatccgagggcccccaagagagcccttcgcgagactacatcaggagcccttccttcagagac 60
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
DB 886 GACATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 945

RESULT 7
AX019132 4307 bp DNA PAT 07-SEP-2000
LOCUS Sequence 2 from Patent W09941397.
DEFINITION AX019132
ACCESSION AX019132
VERSION AX019132.1 GI:10043165
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4307)
AUTHORS Mitrophanous, K., Kingsman, A.J. and Kim, N.
TITLE Anti-viral vectors
JOURNAL Patent: WO 9941397-A 2 19-AUG-1999;
MITROPHANOUS KYPIACOS (GB); KINGSMAN ALAN JOHN (GB); OXFORD
BIOMEDICA LTD (GB); KIM NARRY (KR)
FEATURES
Location/Qualifiers
1. 4307
/organism="Human immunodeficiency virus"
/db_xref="taxon:12721"
BASE COUNT 1114 a 1214 c 1294 g 685 t
ORIGIN
```















GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2001, 20:00:46 ; Search time 5719.75 Seconds  
(without alignments)  
3999.616 Million cell updates/sec

Hit loc: US-09-475-704 - 8  
Perfect score: 1479  
Signature: 1atqqdeeeeqeueqeat.....arqaeceetdaaaagtaa [479]

See the table: IDENTIFY\_NUC  
table 10.0 , (text 1.0)

Searched: 144157 soaps, 773874588 residues

Total number of bits satisfying chosen parameters: 2688314

Maximum I/Os/sec length: 0  
Maximum I/Os/sec length: 2000000000

Post processing: Minimum Match: 0%  
Maximum Match: 100%  
Listed first 45 summaries

```

leafabase :
  1: qb_ba1 : *
  2: qb_ba2 : *
  3: qb_ba3 : *
  4: qb_ba4 : *
  5: qb_ba5 : *
  6: qb_ba6 : *
  7: qb_ba7 : *
  8: qb_ba8 : *
  9: qb_ba9 : *
  10: qb_ba10 : *
  11: qb_ba11 : *
  12: qb_ba12 : *
  13: qb_ba13 : *
  14: qb_ba14 : *
  15: qb_ba15 : *
  16: qb_ba16 : *
  17: qb_ba17 : *
  18: qb_ba18 : *
  19: qb_ba19 : *
  20: qb_ba20 : *
  21: qb_ba21 : *
  22: qb_ba22 : *
  23: qb_ba23 : *
  24: qb_ba24 : *
  25: qb_ba25 : *
  26: qb_ba26 : *
  27: qb_ba27 : *
  28: qb_ba28 : *
  29: qb_ba29 : *
  30: qb_ba30 : *
  31: qb_ba31 : *
  32: qb_ba32 : *
  33: qb_ba33 : *
  34: qb_ba34 : *
  35: qb_ba35 : *
  36: qb_ba36 : *
  37: qb_ba37 : *
  38: qb_ba38 : *
  39: qb_ba39 : *
  40: qb_ba40 : *
  41: qb_ba41 : *
  42: qb_ba42 : *
  43: qb_ba43 : *

```

44:	em_ov: *
45:	em_pat: *
46:	em_ph: *
47:	em_pl: *
48:	em_ro: *
49:	em_sts: *
50:	em_sy: *
51:	em_un: *
52:	em_vl: *
53:	qb_sts1: *
54:	qb_sts2: *
55:	qb_sts3: *
56:	qb_sy: *
57:	qb_un: *
58:	qb_vl: *
59:	qb_vl2: *
60:	qb_htq1: *
61:	qb_htq2: *
62:	qb_htq3: *
63:	qb_htq4: *
64:	qb_htq5: *
65:	qb_htq6: *
66:	qb_htq7: *
67:	qb_htq8: *
68:	qb_htq9: *
69:	qb_htq10: *
70:	qb_htq11: *
71:	qb_htq12: *
72:	qb_htq13: *
73:	qb_htq14: *
74:	qb_htq15: *
75:	qb_htq16: *
76:	qb_htq17: *
77:	qb_htq18: *
78:	qb_htq19: *
79:	qb_htq20: *
80:	qb_htq21: *
81:	qb_htq22: *
82:	qb_htq23: *
83:	qb_htq24: *
84:	qb_htq25: *
85:	qb_pr1: *
86:	qb_pr2: *
87:	qb_pr3: *
88:	qb_pr4: *
89:	qb_pr5: *
90:	qb_pr6: *
91:	qb_pr7: *
92:	qb_pr8: *
93:	qb_pr9: *
94:	qb_pr10: *
95:	qb_pr11: *
96:	qb_pr12: *
97:	qb_pr13: *
98:	qb_pr14: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1197.8	81.0	1509	56	AF201927	Synthetic
2	1159.6	78.4	1548	56	AF287354	Synthetic
3	1114.6	75.4	1847	56	AF202464	Synthetic
4	1114.6	75.4	1847	56	AF202465	Synthetic
5	1091.8	73.8	4352	56	AF287352	Synthetic
6	1066	72.1	4353	56	AF287353	Synthetic
7	949.4	64.2	4307	9	AX035453	Sequence
8	949.4	64.2	4307	9	AX056836	Sequence









```
Db 1441 ATACACAGGAACGTATCCCTTAACCTCCCTCAGATCACTCTTTGGCAACGACCCCTCG 1500
QY 1471 agccagtaa 1479
Db 1501 ICACAGTAA 1509

RESULT 5
AF287352 4352 bp DNA SYN 01-NOV-2000
LOCUS Synthetic construct HIV-1-derived gag protein and pol protein
DEFINITION genes, complete cds.
ACCESSION AF287352
VERSION AF287352.1 GI:11066861
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 4352)
AUTHORS Fuller, M. and Anson, D.S.
TITLE Helper plasmids for production of HIV derived vectors
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4352)
AUTHORS Fuller, M. and Anson, D.S.
TITLE Direct Submission
JOURNAL S. Hmitted (12.01.2000) Chemical Pathology, Women's and Children's
Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES
location/Qualifiers
1..4352
/organism="synthetic construct"
/db_xref="taxon:32630"
/feature="derived from human immunodeficiency virus type 1
isolate YU-2"
37..1539
/note="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="gag protein"
/protein_id="AAG28735.1"
/db_xref="GI:11066862"
/translation="MCAFAVSLACGFLDFWEKTPIMGRKQVPIKHIVWASPELERF
AVPCLLEISECPGIIIGIQLSEIQTCSFEELSYNTVATLCVHOKIEVKDTKEALE
KIEEONKSKKKAQAAADTNSQVSNYPIVONIQQMVHQAISPRTLNAAVKVE
EKAFSEVIMFSALECATPDILNMLNTVGHQAQAMQLKFTINEEAERDLHPV
HAGPTAGQMREPRGSDIAGTSTIQEQIGWMTNPPPIPVGETYKRWIILGLNKIVRM
YSFTSLDIPGPKPEFDDYVDPYKTLPAEQAQSEVKNWMTFTLLVONANPDCKTLL
KALGPAATLEEMTACQGVGGPHKARVLAFAMSQVINSAILMMQRGNFRNKAIVKC
PNCGRSHIAKNCRAPKKGCWKCKCKFGHOMKDCETEROANFLGKIWPCHKRPGNFIO
SRPEPTAPSESVREGERTTTPSQOEPIDKELYPLASRLSLFSGDSPSQ"
1338..4343
/note="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="pol protein"
/protein_id="AAG28736.1"
/db_xref="GI:11066863"
/translation="REDLAFPGQKARKFSFQTRANSPIRRPQVWRDNNLSSEAGA
DPQCTVSFSPQITIQWPPPLVTIKGGQLKEALIDTGDPTVLEEMNIPGRKPKMIG
GIGFTKVRQYDQIPIELICGHAIGVLVGPVPIVNIIGRNLLIQLGCLNFEPSPIET
VPVKLPGMDGPKVKQWPLTEKIKALVEICTEMEKEGKISKIGPENPYNTVFAIKK
KDKTKWPKLIVDFRELNRKLTQDFWEVLGIPHPAGLKKKKSVTIVDGDAYFVFLHED
FRKYTAFTIPISINNETPGTRYQVNVLPQGWKGPSAIFQSSMTTILEPFRKQNDLVLY
QYMDLVVGSDELFIGHRTKIEELKHLIRWGFTTPDKKHQKEPPFLWMGYELHPDKW
TVQPIVLPKDSWTNVNLIQKIVGKLNWASQIYATKVRQICLIRGKALTEVPLTE
EAELEAENREILKEPVHGVYDPSKDLIAEIQOGQGWYQIYQEPFKNLTKGYA
KTRGAHFNIVKQITEAVQKIAFESIVTWKTPKPKIPIQFTWFTWTFYWAQWIPE
WEFVNTPIPLVKLWYQLEKEPIGAETFYVDGAANRETKLKGAGVINKGRQVSLTD
TTNOKTFLQATYIALIDUSGLEVINIVTDSYALGLIQAOQDRSEHLSVQIIFQLIKKE
KVYLAWPFAHKGIGGNEQVDKLVSAIGIRKVLFDLDKDAQEEHEKYHSNWRMASDFN
LPPVAKELIVASCQDKQKGFAMHGOVDCSPGIWOLDCTHLEKGVILVAVHVASGYIE
AEVIPAETQGETAYFLKLAGRWVLIHIDNCSNFIISAIVKAAAWAGIKQEGIPY
NPQSGGVVFSMNKFLKLIQOVRODAEHLKTAQVMAVFIHNFRKGGIGGYSAFPIV
DIIAIDIQIKELQKQITIKQNFVRYKDSNDPLMKGPFAKLLMKGGAVVIQUNSDIKV
```

















647 CTGAAGGAGACATCAATAGGAGAGGCTGCCGAAAGGATGCTGTGATCTGCTGACGCA 706  
 655 gggcccaatcgcccccagatgag 714  
 707 gggcccaatcgcccccagatgag 766  
 715 agcaccctgag 774  
 767 AGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 826  
 775 atctatctatctatctatctatctatctatctatctatctatctatctatctatctatct 834  
 827 AGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 886  
 835 agcaccctgag 894  
 887 AGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 946  
 895 ttccag 954  
 947 TACAAAAGGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1006  
 955 ctgct 1014  
 1007 CTGCT 1066  
 1015 ggcag 1074  
 1067 GGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1126  
 1075 cgcctg 1128  
 1127 GGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1186  
 1129 ggcag 1188  
 1187 GGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1246  
 1189 atcgcccccagatgag 1248  
 1247 AATGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1306  
 1249 cgcctg 1308  
 1307 CAGCAAAATGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1366  
 1309 cgcag 1368  
 1367 CAGCAAAATGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1426  
 1369 ggcag 1416  
 1427 CAGCAAAATGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1486  
 1417 ggcag 1476  
 1487 AAGGAATGCTGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1546  
 1477 taa 1479  
 1547 TAA 1549

RESULT 12  
 AX035463  
 LENGTH 4642 bp  
 DEFINITION Sequence 12 from Patent WO0055341.  
 ACCESSION AX035463  
 VERSION AX035463.1 G1:11191105  
 KEYWORDS  
 SOURCE Synthetic construct  
 ORGANISM

artificial sequence.  
 1 (bases 1 to 4642)  
 Uden, M. and Mitrophanous, K.  
 Anti-viral vectors  
 Patent: WO 0055341-A 12 21-SEP-2000  
 UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) : MITROPHANOUS KYRIACOS (US)  
 Location/Qualifiers  
 1..4642  
 /organism "synthetic construct"  
 /db\_xref "taxon:32630"  
 /note "PSYNEP2 - codon optimized HIV-1 gagpol with leader sequence"  
 BASE COUNT 1218 a 1273 c 1389 g 762 t  
 ORIGIN  
 Query Match 63.5%; Score 959.8; Db 9; Length 4642;  
 Best Local Similarity 78.6%; Pred. No. 150-97;  
 Matches 1182; Conservative 0; Mismatches 297; Indels 24; Gaps 4;  
 QY 1 atgggc 60  
 DB 336 ATGGGCG 395  
 QY 61 ctggagc 120  
 DB 396 CTGGAGCG 455  
 QY 121 ctggagc 180  
 DB 456 CTGGAGCG 515  
 QY 181 atcgcccccagatgag 240  
 DB 516 ATCGCCCG 575  
 QY 241 accctgagc 300  
 DB 576 ACCGTGGCG 635  
 QY 301 ctggagc 354  
 DB 636 CTGGAGCG 695  
 QY 355 ggc 414  
 DB 696 GATGAGCG 755  
 QY 415 cgcctgagc 474  
 DB 756 CAGATGAGCG 815  
 QY 475 ggc 534  
 DB 816 GAGGAGCG 875  
 QY 535 cgc 594  
 DB 876 CCGCAAGATGCTGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 935  
 QY 595 ctggagc 654  
 DB 936 CTGGAGCG 995  
 QY 655 ggc 714  
 DB 996 GGGCCGATCG 1055  
 QY 715 agcaccctgag 774  
 DB 1056 AGTACGCTTCAGGAAAGATCGGCTGGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1115















Query Match	Score	DB	Length
Query Match	53.9%	Score 797.4	DB 58
Best Local Similarity	71.2%	Prod. No. 16-81	
Query Match	53.9%	Score 797.4	DB 58
Best Local Similarity	71.2%	Prod. No. 16-81	



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 8, 2001, 29:59:23 ; Search time 3475.48 Seconds  
(without alignments)  
4104,284 Million cell updates/sec

Title: US 04-475-704-4  
 Perfect Score: 1509

Sequence: 1 atqqqqccccqccagat ..... qqqqccctgagcactaa 1509

Scoring table: `IDENTITY_NUC`  
Gapop 10.0 ,

Searches: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456280

```
Minimum DB Seq length: 0
Maximum DB Seq length: 2000000000
```

Post-processing: Minimum Match 9%  
Maximum Match 100%  
Listing first 45 summaries

```

Database :
EST: *
1: qb_est1: *
2: qb_est2: *
3: qb_est3: *
4: qb_est4: *
5: qb_est5: *
6: qb_est6: *
7: qb_est7: *
8: qb_est8: *
9: qb_est9: *
10: qb_est10: *
11: qb_est11: *
12: qb_est12: *
13: qb_est13: *
14: qb_est14: *
15: qb_est15: *
16: qb_est16: *
17: qb_est17: *
18: qb_est18: *
19: qb_est19: *
20: qb_est20: *
21: qb_est21: *
22: qb_est22: *
23: qb_est23: *
24: qb_est24: *
25: qb_est25: *
26: qb_est26: *
27: qb_est27: *
28: qb_est28: *
29: qb_est29: *
30: qb_est30: *
31: qb_est31: *
32: qb_est32: *
33: em_est1a: *
34: em_est1a: *
35: em_est1a: *
36: em_est1a: *
37: em_est1a: *
38: em_est1a: *
39: em_est1a: *
40: em_est1a: *
41: em_est1a: *
42: em_est1a: *
43: em_est1a: *

```

117: qb\_est48: \*  
118: qb\_est49: \*  
119: qb\_est50: \*  
120: qb\_est51: \*  
121: qb\_est52: \*  
122: qb\_est53: \*  
123: qb\_est54: \*  
124: qb\_est55: \*  
125: qb\_est56: \*  
126: qb\_est57: \*  
127: qb\_est58: \*  
128: qb\_est59: \*  
129: qb\_est60: \*  
130: qb\_est61: \*  
131: qb\_est62: \*  
132: qb\_est63: \*  
133: qb\_est64: \*  
134: qb\_est65: \*  
135: qb\_est66: \*  
136: qb\_est67: \*  
137: qb\_est68: \*  
138: qb\_est69: \*  
139: qb\_est70: \*  
140: qb\_est71: \*  
141: qb\_est72: \*  
142: qb\_est73: \*  
143: qb\_est74: \*  
144: qb\_est75: \*  
145: qb\_est76: \*  
146: qb\_est77: \*  
147: qb\_est78: \*  
148: qb\_est79: \*  
149: qb\_est80: \*  
150: qb\_est81: \*  
151: qb\_est82: \*  
152: qb\_est83: \*  
153: qb\_est84: \*  
154: qb\_est85: \*  
155: qb\_est86: \*  
156: qb\_est87: \*  
157: qb\_est88: \*  
158: qb\_est89: \*  
159: qb\_est90: \*  
160: qb\_est91: \*  
161: qb\_est92: \*  
162: qb\_est93: \*  
163: qb\_est94: \*  
164: qb\_est95: \*  
165: qb\_est96: \*  
166: qb\_est97: \*  
167: qb\_est98: \*  
168: qb\_est99: \*  
169: qb\_est100: \*  
170: qb\_est101: \*  
171: qb\_est102: \*  
172: qb\_est103: \*  
173: qb\_est104: \*  
174: qb\_est105: \*  
175: qb\_est106: \*  
176: qb\_est107: \*  
177: qb\_est108: \*  
178: qb\_est109: \*  
179: qb\_est110: \*  
180: qb\_est111: \*  
181: qb\_est112: \*  
182: qb\_est113: \*  
183: qb\_est114: \*  
184: qb\_est115: \*  
185: qb\_est116: \*  
186: qb\_est117: \*  
187: qb\_est118: \*  
188: qb\_est119: \*  
189: qb\_est120: \*

190: qb\_est121: \*  
191: qb\_est122: \*  
192: qb\_est123: \*  
193: qb\_est124: \*  
194: qb\_est125: \*  
195: qb\_est126: \*  
196: qb\_est127: \*  
197: qb\_est128: \*  
198: qb\_est129: \*  
199: qb\_est130: \*  
200: qb\_est131: \*  
201: qb\_est132: \*  
202: qb\_est133: \*  
203: qb\_est134: \*  
204: qb\_est135: \*  
205: qb\_est136: \*  
206: qb\_est137: \*  
207: qb\_est138: \*  
208: qb\_est139: \*  
209: qb\_est140: \*  
210: qb\_est141: \*  
211: qb\_est142: \*  
212: qb\_est143: \*  
213: qb\_est144: \*  
214: qb\_est145: \*  
215: qb\_est146: \*  
216: qb\_est147: \*  
217: qb\_est148: \*  
218: qb\_est149: \*  
219: qb\_est150: \*  
220: qb\_est151: \*  
221: qb\_est152: \*  
222: qb\_est153: \*  
223: qb\_est154: \*  
224: qb\_est155: \*  
225: qb\_est156: \*  
226: qb\_est157: \*  
227: qb\_est158: \*  
228: qb\_est159: \*  
229: qb\_est160: \*  
230: qb\_est161: \*  
231: qb\_est162: \*  
232: qb\_est163: \*  
233: qb\_est164: \*  
234: qb\_est165: \*  
235: qb\_est166: \*  
236: qb\_est167: \*  
237: qb\_est168: \*  
238: qb\_est169: \*  
239: qb\_est170: \*  
240: qb\_est171: \*  
241: qb\_est172: \*  
242: qb\_est173: \*  
243: qb\_est174: \*  
244: qb\_est175: \*  
245: qb\_est176: \*  
246: qb\_est177: \*  
247: qb\_est178: \*  
248: qb\_est179: \*  
249: qb\_est180: \*  
250: qb\_est181: \*  
251: qb\_est182: \*  
252: qb\_est183: \*  
253: qb\_est184: \*  
254: qb\_est185: \*  
255: qb\_est186: \*  
256: qb\_est187: \*  
257: qb\_est188: \*  
258: qb\_est189: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

































117: qb\_est48: \*  
118: qb\_est49: \*  
119: qb\_est50: \*  
120: qb\_est51: \*  
121: qb\_est52: \*  
122: qb\_est53: \*  
123: qb\_est54: \*  
124: qb\_est55: \*  
125: qb\_est56: \*  
126: qb\_est57: \*  
127: qb\_est58: \*  
128: qb\_est59: \*  
129: qb\_est60: \*  
130: qb\_est61: \*  
131: qb\_est62: \*  
132: qb\_est63: \*  
133: qb\_est64: \*  
134: qb\_est65: \*  
135: qb\_est66: \*  
136: qb\_est67: \*  
137: qb\_est68: \*  
138: qb\_est69: \*  
139: qb\_est70: \*  
140: qb\_est71: \*  
141: qb\_est72: \*  
142: qb\_est73: \*  
143: qb\_est74: \*  
144: qb\_est75: \*  
145: qb\_est76: \*  
146: qb\_est77: \*  
147: qb\_est78: \*  
148: qb\_est79: \*  
149: qb\_est80: \*  
150: qb\_est81: \*  
151: qb\_est82: \*  
152: qb\_est83: \*  
153: qb\_est84: \*  
154: qb\_est85: \*  
155: qb\_est86: \*  
156: qb\_est87: \*  
157: qb\_est88: \*  
158: qb\_est89: \*  
159: qb\_est90: \*  
160: qb\_est91: \*  
161: qb\_est92: \*  
162: qb\_est93: \*  
163: qb\_est94: \*  
164: qb\_est95: \*  
165: qb\_est96: \*  
166: qb\_est97: \*  
167: qb\_est98: \*  
168: qb\_est99: \*  
169: qb\_est100: \*  
170: qb\_est101: \*  
171: qb\_est102: \*  
172: qb\_est103: \*  
173: qb\_est104: \*  
174: qb\_est105: \*  
175: qb\_est106: \*  
176: qb\_est107: \*  
177: qb\_est108: \*  
178: qb\_est109: \*  
179: qb\_est110: \*  
180: qb\_est111: \*  
181: qb\_est112: \*  
182: qb\_est113: \*  
183: qb\_est114: \*  
184: qb\_est115: \*  
185: qb\_est116: \*  
186: qb\_est117: \*  
187: qb\_est118: \*  
188: qb\_est119: \*  
189: qb\_est120: \*

190: qb\_est121: \*  
191: qb\_est122: \*  
192: qb\_est123: \*  
193: qb\_est124: \*  
194: qb\_est125: \*  
195: qb\_est126: \*  
196: qb\_est127: \*  
197: qb\_est128: \*  
198: qb\_est129: \*  
199: qb\_est130: \*  
200: qb\_est131: \*  
201: qb\_est132: \*  
202: qb\_est133: \*  
203: qb\_est134: \*  
204: qb\_est135: \*  
205: qb\_est136: \*  
206: qb\_est137: \*  
207: qb\_est138: \*  
208: qb\_est139: \*  
209: qb\_est140: \*  
210: qb\_est141: \*  
211: qb\_est142: \*  
212: qb\_est143: \*  
213: qb\_est144: \*  
214: qb\_est145: \*  
215: qb\_est146: \*  
216: qb\_est147: \*  
217: qb\_est148: \*  
218: qb\_est149: \*  
219: qb\_est150: \*  
220: qb\_est151: \*  
221: qb\_est152: \*  
222: qb\_est153: \*  
223: qb\_est154: \*  
224: qb\_est155: \*  
225: qb\_est156: \*  
226: qb\_est157: \*  
227: qb\_est158: \*  
228: qb\_est159: \*  
229: qb\_est160: \*  
230: qb\_est161: \*  
231: qb\_est162: \*  
232: qb\_est163: \*  
233: qb\_est164: \*  
234: qb\_est165: \*  
235: qb\_est166: \*  
236: qb\_est167: \*  
237: qb\_est168: \*  
238: qb\_est169: \*  
239: qb\_est170: \*  
240: qb\_est171: \*  
241: qb\_est172: \*  
242: qb\_est173: \*  
243: qb\_est174: \*  
244: qb\_est175: \*  
245: qb\_est176: \*  
246: qb\_est177: \*  
247: qb\_est178: \*  
248: qb\_est179: \*  
249: qb\_est180: \*  
250: qb\_est181: \*  
251: qb\_est182: \*  
252: qb\_est183: \*  
253: qb\_est184: \*  
254: qb\_est185: \*  
255: qb\_est186: \*  
256: qb\_est187: \*  
257: qb\_est188: \*  
258: qb\_est189: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.







































Db	1423	TCACAGAAAGAGAGGAGATACAAAGAACTGTAATCTTTAACTTCCTCAGATCACTC	1482
QY	1483	ttcggcagcggccctcctgagccagtaa	1509
Db	1483	TTTGGCAACGACCCCTCGTCACAGTAA	1509
RESULT	5		
AF287352			
LOCUS	AF287352	4352 bp	DNA
DEFINITION	Synthetic construct HIV-1-derived gag protein and pol protein genes, complete cds	SYN	01-NOV-2000
ACCESSION	AF287352		
VERSION	AF287352.1	GI:11066861	
KEYWORDS			
SOURCE	Synthetic construct.		
ORGANISM	Synthetic construct		
REFERENCE	1 (bases 1 to 4352)		
AUTHORS	Fuller, M. and Anson, D.S.		
TITLE	Helper plasmids for production of HIV derived vectors		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 4352)		
AUTHORS	Fuller, M. and Anson, D.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-JUL-2000) Chemical Pathology, Women's and Children's Hospital, 72 King William Road, North Adelaide, SA 5006, Australia		
FEATURES	Location/Qualifiers		
source	1..4352		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="derived from Human immunodeficiency virus type 1 isolate YU-2"		
	37..1539		
	/note="codon optimized reading frame"		
	/codon_start=1		
	/transl_table=11		
	/product="gag protein"		
	/protein_id="AAG28735.1"		
	/db_xref="GI:11066862"		
	/translation="MCAASVLSAGELDKWEKIFLPSSKKQYPIKHIWASPFIFRF		
	AVPDSIFTSERFQILGQLPSLQTSSELPISLYNTVATLYCVHOKIEVKDTKEALE		
	KIFEQNKSKKKAQAAADTGNSSQVSNYPVQNLQGMVHQAIISPTLNAMVKVVE		
	EKAESPEVIMPSALSAGATPQDINTMLNTVGGHQAAMQMLKETINEEAEDRLHPV		
	HAGPIAGQMPPFRGSDIAGTTSTLOEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM		
	YSPTSLDIRQGPKEPRDYVDFYKTIKRAQASQEVKNMTETLLVQANAPDCKTIL		
	KALGPAATLEEMFIACGVGGPGHKARVLALAMSQVINSALIMMHPGFNFNFKTVKC		
	FNCKECHIAKNTPAPPKKGCKGCKGCKGCKGCKGCKGCKGCKGCKGCKGCKGCKG		
	SRPEPTAPSESVRFGEETTPSQKQEPIDKELYPLASLSRSLFGSDPSSQ"		
	<1338..4343		
	/note="codon optimized reading frame"		
	/codon_start=1		
	/transl_table=11		
	/product="pol protein"		
	/protein_id="AAG28736.1"		
	/db_xref="GI:11066863"		
	/translation="REDLAFPOGKARKFSSEQTPANSPIHFFPQVWFFPNNNSI SEAGA		
	PRQGVVSFPQITLWQPLVITKIGGOLKEALLDGTADDTVLEEMNIPGRWPKMIG		
	GIGGFIKVRQYDQIPIEICGHKAIGTLVGPVFNILGKLLTQIGCILNFPISPIET		
	VPVKIKPGMDGPKVKQWPLTEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKK		
	KDSTKPKLVDFPELNKPTQDFEVLGIGIPHPAGLKKKSVTLVDVGDAVFSVPLHED		
	EPFYTAFTIPSNINTPGTGYQVNLVPGWKGPSAIPQSSMTTILEPFRKQNPDLVY		
	QYMDLLVGSDELIGQHRTKIEELRQHLLRWGFTTPDKKHOKEPPELWMSYELHPDKW		
	TVQPIVLPFKDSWTNDIQIKLVKNWASQIYAGIKVRQILCKLLRGTKALTEVILPTE		
	EAELELAENREILKEPVHGVYDPSKDLIAELIKQGGQWLYIYQEPFNKIKGYA		
	PTGATNDVVKQITAVQKTIATESIVTWGKTPKFLPIQKETWETWTWTEYQWATWIPE		
	WEFVNTPLVLKWLQLEKEPIIGATFYVDGAANKETIKLGKAGYVINKGKQKVVSLID		
	TTNQKTELQAIYALQDSGLEVINVTDSQYALGIQAQPDSESELSVQIIEQLIKKE		
	KVYLAWVPAHKGIGGNEVDKIVSAGIRKVLFDGIDKQAEHFKYHSNWRAMSDFN		
	LPPVVAKEIVASCKQIKGFAMHGGVDCSPGTLQDLNTHLPKQVILVAHVASGYIE		
	AEVTPARTQETAYEFLKLAGRPVVIHLUNGSNFISAIKAAACWAGIKGEFGIPY		
	NPQSGVVFESMKNKFLKFTIGVPRQAFHKTAVQMAVFIHNFKRKGGIGGYSAGERIV		
	DLIATDIQTKELQKQITKIQNFPVYVPSRDPDLWKGPALILWKGRGAVVIQDNSDIKV		















Db 381 GACACGGGACACAGCAACAGGATCAGCCAGAACTACCCCATCGTCCAGAACATCCAGGGG 440  
QY 412 cagatggtgcacagggcagcagcccccagacccctgaacccctggtatgaagatcagcag 471  
Db 441 CAGATGGTGCACAGGCCATCTCCCGCCGACCGCTGAAGCCCTGCTGAAGGTGGTGA 500  
QY 472 gagaagcccttcagcccccaggtgatcccatgttccacccctgaacagagggccacc 531  
Db 501 GAGAAAGGCTTTTACAGCCGAGGTGATACATGCTTCTCAGCCCTGCTGAAGGGAGCCACC 560  
QY 532 ccccccagccctgaacacagatgttgaaacacccgtggcggcagccagggccctgaacatg 591  
Db 561 CCCCAGAGATCTGAACATGCTCAACACAGTGGGGGACACCAAGCGCCCATCCAGATG 620  
QY 592 ctgagggacacacatcaacagagggcggcggagtggaacccctgcaccccttcacagcc 651  
Db 621 CTGAAGGAGACCATCAATGAGGAGGCTGCCGAATGGGATCGTGTGCTATCCGGTGCACGCA 680  
QY 652 ggcgccgtggcccccggcagatggcggcagcccccggcggcagccacatcccccagcc 711  
Db 681 GGGGCTATGCAAGCGCCAGATGCTGAGCAACGGGCTCAGACATCGCGGGAACGACI 740  
QY 712 ggcacccctgcagggagcagcagcggcggcggcggcggcggcggcggcggcggcggcggc 771  
Db 741 AGTACCTTCAGGAACAGATCGGCTGGATGACCAACACACCCATCCATCCGGTGGGAGAA 800  
QY 772 atctcaagcgggtgagatccctggcggcggcggcggcggcggcggcggcggcggcggc 831  
Db 801 ATCTACAAACGCTGGATCATCTGAGGCTGAACAAAGATCGTGGCATGTATAGCCCTACC 860  
QY 832 agcatccctgcagcagcagcggcggcggcggcggcggcggcggcggcggcggcggcggc 891  
Db 861 AGTATCTGAGCAATCTGAGTGAAGGCTGAGCAACCTTTCTGCAATGCTGCTGCTT 920  
QY 892 ttcagagccctgcgcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggc 951  
Db 921 TACAAACGCTCCCGCGGAGCAGCGCTAGCCAGGAGGTGAAGAACTGGATGACCGAAACC 980  
QY 952 ctgctggtgcagaaacgcaaccccgagcgaagaccatccctgcgcggcggcggcggcggc 1011  
Db 981 CTGCTGGTTCAGAAACCGGACCCGGACTGCAAGAGGATCCTGAAGGCCCTGGGCCCCAGGG 1040  
QY 1012 gccacccctggagggatgatgacccctgcagcggcggcggcggcggcggcggcggcggc 1071  
Db 1041 GCTACCCCTAGAGGAAATGATGACCGCTGTCAGGAGTGGCGGAGCCGCGCACAAAGGCA 1100  
QY 1072 cgcctgctgagcggagggatgagcagg--ccaaacagcgtgaacatcatgatcagagag 1128  
Db 1101 CAGGCTTGAAGGCTGAGCCAGTACCCAGTGAAGTGGCGGAGCCGCGCACAAAGGCA 1160  
QY 1129 agcaacttcaagggcccccgccgcaagtcgaagtcttcaactgctgagcggcggcggcggc 1188  
Db 1161 GGCAACTTTGGGAAACCAACCCAGATCGTCAAGTGTCTCAACTGAGCAAAAGAGGCA 1220  
QY 1189 atgcgcaagaaactgcccgcggcccccgcaagaaggcgtgctggaatgctgagcggcggc 1248  
Db 1221 ACAGCCCGGAAATGAGGGGCTCTAGGAAAGGGCTGTGGAAATGTGGAAAGGAAAGGA 1280  
QY 1249 caccagatgaaggaactgcacccagcgcggcggcggcggcggcggcggcggcggcggcggc 1308  
Db 1281 CACCAATGAAGATTTGATGAGAGACAGGCTAAATTTTATGGGAAGATCTGGGCTTCC 1340  
QY 1309 cacaagggccgccccggcgaacttctggaacccgcagcagcggcggcggcggcggcggcggc 1368  
Db 1341 CACAAGGGAAGCCAGGGAATTTCTTCAGAGCAGACACAGAGC----- 1383  
QY 1369 cccacccgcccccccgccagcgttcccttcagagagaccccccagcggcggcggcggcggc 1428  
Db 1384 -CAACAGCCCCACAGAGAGAGCTCAGGTTTGGGAAAGAGACAAACCTCCCTCTAG 1442  
QY 1429 gaccccaaggaacgcggagcctacccgagagccctgaccccccctgcgcagcctgttcagg 1488  
Db 1443 AAGCAGGAGCCGATAGACAAAGGAACIGTATCTTTAGCTTCCCTCAGATCACTCTTCC 1502

QY 1489 agcgcgcgcctgagccagtaa 1509  
Db 1503 AGGAGCCCTCGTCACAATAA 1523  
RESULTS 11  
AX035464 4353 bp DNA PAT 15-NOV-2000  
LOCUS Sequence 13 from Patent WO0055341  
DEFINITION AX035464  
ACCESSION AX035464  
VERSION AX035464.1 GI:1191186  
KEYWORDS synthetic construct,  
SYNOPSIS synthetic construct,  
ORGANISM artificial sequence,  
REFERENCE 1 (bases 1 to 4353)  
AUTHORS Uden,M. and Mitrophanous,K.  
TITLE Anti-viral vectors  
JOURNAL Patent: WO 0055341-A 13 21-SEP-2000;  
UDEN MAPK (GB) ; OXFORD BIOMEDICA LTD (GB) ; MITROPHANOUS KYRIACOS  
(US)  
FEATURES  
Location/Qualifiers  
source 1..4353  
/organism="synthetic construct"  
/db\_xref="taxon:4630"  
/note="psYNGP3 - codon optimised HIV-1 gagpol with leader  
sequence from the major splice donor"  
BASE COUNT 1153 a 1198 c 1302 g 700 t  
ORIGIN  
Query Match 61 38. Score 925.4; DB 9; Length 4353;  
Best Local Similarity 77.36, E-Val 1.5e-92;  
Matches 1175; Conservative 0; Mismatches 316; Indels 30; Gaps 3;  
QY 1 atggggc 60  
Db 47 ATGGGCG 106  
QY 61 ctggc 120  
Db 107 CTGGGCG 166  
QY 121 ctggaggggttgc 180  
Db 167 CTGGGCG 226  
QY 181 atgagcagc 240  
Db 227 CTGGGCG 286  
QY 241 accgtggc 300  
Db 287 AGGCG 346  
QY 301 ctggcagagatcggagggagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360  
Db 347 CTGGGCG 406  
QY 361 ggc 411  
Db 407 GACACCGGACACAGCAACCCAGGTGAGCGAGCAACTACCCCATCTCTCGCAAGCATCCAGGGG 466  
QY 412 cagatggtgcacccagccatcagcccccgccacccctgaacccctgagcagcagcagcagcagc 471  
Db 467 CAGATGGTGCACAGGCCATCTCCCGCCGACCGCTGCAACCGCTGCGGTGAAGGTGGTGGAA 526  
QY 472 gagaagcccttcagcccccaggtgatcccatgttccacccctgaacagcagcagcagcagcagc 531  
Db 527 GAGAAAGGCTTTTACAGCCGAGGTGATACATGCTTCTCAGCCCTGCTCAGAGGAGCCACC 586  
QY 532 cccagagcctgaacagatgttgaaacccgtggggggcgaacccagccagccagcagcagcagc 591















DE 1717 AGGCAAGCGTACAGGGAACCGTTAACTGGCGCTCACATCACTCTTTGGCAGCGGCGCGTTG 1776

QY 1501 aqccagtaa 1509  
|||

Db 1777 TCTCAATAA 1785

RESULT 15

AF286234

LOCUS AF286234 9034 bp DNA VRL 10-APR-2001

DEFINITION HIV-1 strain 98T2013 from Tanzania, complete genome.

ACCESSION AF286234

VERSION AF286234.1 GI:13569317

KEYWORDS

SOURCE Human immunodeficiency virus type 1.

ORGANISM Human immunodeficiency virus type 1

REFERENCE: 1 (bases 1 to 9034)

AUTHORS Podenburg, C.M., Li, Y., Trask, S.A., Chen, Y., Decker, J., Robertson, D.L., Kalish, M.L., Shaw, G.M., Allen, S., Hahn, B.H. and Gao, F.

TITLE Near full-length clones and reference sequences for subtype C isolates of HIV type 1 from three different continents

JOURNAL AIDS Res. Hum. Retroviruses 17 (2), 161-168 (2001)

MEDLINE 21094715

REFERENCE 2 (bases 1 to 9034)

AUTHORS Rodenburg, C.M., Li, Y., Trask, S.A., Chen, Y., Decker, J., Robertson, D.L., Allen, S., Shaw, G.M., Hahn, B.H. and Gao, F.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-2000) Medicine, University of Alabama at Birmingham, 701 19th Street, South, LHRB 641, Birmingham, AL 35294, USA

FEATURES

source location/Qualifiers

1..9034

/organism="Human immunodeficiency virus type 1"

/proviral

/strain="98T2013"

/db\_xref="taxon:11676"

/country="Tanzania"

/note="isolated in 1998"

136..1635

/gene="gag"

136..1635

/gene="gag"

/codon\_start=1

/product="gag protein"

/protein\_id="AAK31053.1"

/db\_xref="GI:13569318"

/translation="MGARASILKGGKLDKWEIIRLRPGGKKHYMIKHLVWASRELERF  
ALNPSLETADGGKQIIRQLHPALQTGTEELRSLFNTVATLYCVHAEIDVDRDTKEALD  
KIEEONKSOQKTQAKAADVKISQNPYIVQNLKGMQVHKQAIISPTLNANWVYIEEKA  
FSPEEIVPMFATLSGATPDQDLNTMLNTVGGHQAAMQMLKDTINEAAEWDRLHPVHAG  
PIAPGQMKRPGSDIAGTISILQEQIAWMTGNPPVPGVDIYKRWIILGLNKIVRMYSF  
VSIIDIRQCPKEPRDYVDRFFKTLRAEQATODVKNWMTDTLLVQNANPDCKTILRAL  
GPAATLEEMTACQGVGGPSHKARVLAEMSQTSTNIMQRSNEFGPKRTVKCFNCG  
KEGHIARNCRAPKKGCGWKCGRGHQMKGCTERCANFLGKIWPSHKGRPGNPLQGRPE  
PTAPPAEPTAPPAESFRFEGTTTAPKQEPKDRPLETSLKSLFGNDPLFQ"

<1419..4439

/gene="pol"

<1419..4864

/gene="pol"

/codon\_start=1

/product="truncated pol polyprotein"

/protein\_id="AAK31054.1"

/db\_xref="GI:13569319"

/translation="FFPFNIAFPQGEAPKFPSSGOARANSPTSRANSPTSREILOVRGDIN  
PRSETGARERQGTILNPQIALWQRPPLYSIKVGGQVKEALLDTGADPTVLEEINLPQKWK  
PKMIGIGGFIKVRQYDQIPIFICGKKALGTVLGPTPVNIIGRNMLTQLGCTLNFPPI  
SPIETVPVKLPGMDGPKVQWPLTEEKIKALTAICEEMEKEGKITIKIGPENPYNTPV  
FAIKKDKSTKWRKLVDRELNKTDQFWEVQIGTIPHPAGLKKKKSIVILVDVGDAEYFS  
PLDEGFRKYTAFTIPSTNNETPGIRYOYNVLPOGMKGSPIAFQSMTKILEPFIQNP  
GIVIVYOYDDLLVYGSDLIEIGOHRAKVEELREHLIKWGFTTPDKKHOKEPFPLMWGYEL

gene

CDS















































